UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/982,269	10/17/2001	Benoit Mory	PHFR 000110	7787

7590 10/07/2004

US PHILIPS CORPORATION
INTELLECTUAL PROPERTY DEPART
580 WHITE PLAINS ROAD
TARRYTOWN, NY 10591

EXAMINER	
STEVENS, ROBERT	
ART UNIT	**PAPER NUMBER**
2176	

DATE MAILED: 10/07/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/982,269	MORY ET AL.	
	Examiner	Art Unit	
	Robert M Stevens	2176	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *17 October 2001*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-10* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-10* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *17 October 2001* is/are: a)☐ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☒ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *10/17/01, 2/27/02*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

1. Claims 1-10 are pending in Application No. 09/982,269, entitled "Binary Format for MPEG-7 Instances", filed 10/17/2001 by Mory et al. Claims 1, 3, and 5-10 are independent.

2. The Office acknowledges two Information Disclosure Statements filed on 10/17/2001 and 2/27/2002.

3. Acknowledgment is made of applicant's claim for foreign priority based on an application filed in the European Patent Office (EPO) on Oct. 17, 2000. It is noted, however, that applicant has not filed a certified copy of the EP 00402876.7 application as required by 35 U.S.C. 119(b).

Drawings

4. Regarding Fig. 1, 5 and 6: No reference characters (refer to 37 CFR 1.84(p)) appear in these drawings and the associated specification. Reference characters are required to understand the Application subject matter.

5. Regarding Fig. 4, the steps discussed in the specification do not appear in the figure.

6. Figure 2 contains references to steps 2-1 thru 2-4. However, no lead lines appear from the steps to the referenced elements, as required by 37 CFR 1.84(q). Additionally, reference characters are not to be encircled, as per 37 CFR 1.84(p)(1).

7. Figure 3 contains steps 3-1, 3-2 and 3-4. "Step 3-4" appears to be an error. Additionally, no lead lines appear connecting the reference step numbers to the figure elements. Additionally, the empty boxes associated with the Fig. 3 steps require suitable legends, as per 37 CFR 1.84(o). Additionally, the underlining of reference characters is generally associated with a cross section (see 37 CFR 1.84(p)(3)), which does not appear to be Applicant's intent for Fig. 3.

8. Corrected drawing sheets are required in reply to the Office action to avoid abandonment of the application. Any amended replacement drawing sheet should include all of the figures appearing on the immediate prior version of the sheet, even if only one figure is being amended. The figure or figure number of an amended drawing should not be labeled as "amended." If a drawing figure is to be canceled, the appropriate figure must be removed from the replacement sheet, and where necessary, the remaining figures must be renumbered and appropriate changes made to the brief description of the several views of the drawings for consistency. Additional replacement sheets may be necessary to show the renumbering of the remaining figures. The replacement sheet(s) should be labeled "Replacement Sheet" in the page header (as

per 37 CFR 1.84(c)) so as not to obstruct any portion of the drawing figures. If the changes are not accepted by the examiner, the applicant will be notified and informed of any required corrective action in the next Office action. The objection to the drawings will not be held in abeyance.

Specification

9. The abstract of the disclosure is objected to because it is not limited to a single paragraph of 50 – 150 words, and contains claim language. Correction is required. See MPEP § 608.01(b).

10. The disclosure is difficult to understand because multiple terms describe the same element in the figures (e.g., XML instance, XML hierarchy, hierarchical XML structure, XML instance XML-D, instance XML-C). This is the main reason why reference characters are required for the drawings (and the associated description within the specification). Proper use of reference characters ensures consistency in identification of drawing elements.

11. The disclosure is objected to because of the following informalities:

A. Page 12 line 22 references a "Step 2", which is not in the drawing; Applicant is reminded to please correct all spelling/grammatical/etc. mistakes throughout the specification (including the claims and drawings);

B. Page 4 line 33: An inferencing mechanism is alluded to, but not described here (see "inferred"). Is such a mechanism well known in the art? Please explain.

Appropriate correction is required.

Claim Rejections - 35 USC § 101

12. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

13. **Claims 9-10 are rejected under 35 U.S.C. 101** for the following reasons:

Regarding independent claim 9, a "signal" is not tangibly embodied and its usefulness is unclear.

Regarding independent claim 10, the claim is to a "table", which is a software artifact that is not tangibly embodied.

Claim Rejections - 35 USC § 112

14. The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

15. **Claims 1-10 are rejected under 35 U.S.C. 112, second paragraph**, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claims 1-10 are vague and indefinite because these claims, either directly or indirectly via claim dependency, use the term "XML-like", which renders the scope indeterminable.

Claims 1-10 are also vague and indefinite because these claims, either directly or indirectly via claim dependency, use the pronoun "it", which renders the scope indeterminable.

Further regarding claim 2, there is a lack of antecedent basis for "A coding method as claimed in claim1".

Claim 9 recites an intended use for a signal, but the claim language does not set forth any limitations on such signal. This renders the scope of the claim indeterminable.

Claim 10 recites an intended use for a table, but the claim language does not set forth any limitations on such table. This renders the scope of the claim indeterminable.

Claim Rejections - 35 USC § 102

16. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

17. **Claims 9 and 10 are rejected under 35 U.S.C. 102(b)** as being anticipated by Gwendal Auffret, et al., (paper entitled: "Audiovisual-based Hypermedia Authoring: Using Structured Representations for Efficient Access to AV Documents", Hypertext '99, Darmstadt, Germany, Feb. 1999, hereafter referred to as "Auffret").

Regarding independent claim 9, Auffret discloses:

A signal for transmission over a transmission network comprising an encoder and/or a decoder (p. 175 "Structure encoding using XML") *having a memory storing at least one table derived from an XML-like schema* (Fig. 11), *said XML-like schema defining a hierarchical structure of description elements, said hierarchical structure comprising hierarchical levels, parent description elements and child description elements* (p. 174 Fig. 7, and first paragraph under "Temporal Model" re: "graph containing description object"), *said table containing identification information for solely identifying each description element in a hierarchical level* (p. 174 Fig. 7, and first paragraph under "Temporal Model" re: "graph containing description object"), *and structural information for retrieving any child description element from its parent description element* (p. 174 Fig. 7, and first paragraph under "Temporal Model" re: "reference links [structural information]"), *said signal embodying at least one fragment representing a content of a description element* (p. 174 "A segment" section, which also references Fig. 4, showing how sequenced segments are used in the building of a document), *called encoded description element, and a sequence of identification information being associated in said table to said encoded description element and its parent description element(s)* (p. 173 Fig. 4, and p. 173 Fig. 4 re: the last paragraph before the section entitled "Relating Descriptors to an Ontology" and discussing tree building).

Regarding independent claim 10, Auffret discloses:

A table (Fig. 11) *intended to be used in an encoder for encoding a description element of an instance of an XML-like schema, and/or in a decoder for updating a hierarchical memory representation of an instance of an XML-like schema* (p. 175 "Structure Encoding using XML"),

said XML-like schema defining a hierarchical structure of description elements, said hierarchical structure comprising hierarchical levels, parent description elements and child description elements, characterized in that it is derived from said XML-like schema (p. 174 Fig. 7, and p. 173 first sentence under heading "Overview of AEDI"),

and it contains identification information for solely identifying each description element in a hierarchical level (p. 174, Fig. 7 and first paragraph under heading "Temporal Model", re: "graph containing description objects"),

and structural information for retrieving any child description element from its parent description element. (p. 174, Fig. 7 and first paragraph under heading "Temporal Model", re: "reference links [i.e., structural information]")

Claim Rejections - 35 USC § 103

18. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

19. **Claims 1, 3 and 5-8 are rejected under 35 U.S.C. 103(a)** as being unpatentable over Gwendal Auffret, et al., (paper entitled: "Audiovisual-based Hypermedia Authoring: Using Structured Representations for Efficient Access to AV Documents", Hypertext '99, Darmstadt, Germany, Feb. 1999, hereafter referred to as "Auffret") in view of Simon

North, et al., (SAMS Teach Yourself XML in 21 Days, Sam's Publishing, Indianapolis, IN, (c) 1999, hereafter referred to as "North").

Regarding independent method claim 1, Auffret discloses:

A encoding method for encoding a description element of an instance of an XML-like schema defining a hierarchical structure of description elements (p. 175 "Structure encoding using XML"), *said hierarchical structure comprising hierarchical levels, parent description elements and child description elements* (p. 174, Fig. 7 and 1st paragraph under heading "Temporal Model" re: "graph containing description objects"), *said description element to be encoded comprising a content* (p. 174, Fig. 7 and 1st paragraph under heading "Temporal Model" re: "reference links [structural information]"), *characterized in that it consists in:*

> *using at least one table derived from said schema* (Fig. 11), *said table containing identification information for solely identifying each description element in a hierarchical level, and structural information for retrieving any child description element from its parent description element* elements (p. 174, Fig. 7 and 1st paragraph under heading "Temporal Model" re: "graph containing description objects"),
>
> *encoding said description element to be encoded as a fragment comprising said content and a sequence of the retrieved identification information.* content (p. 174, Fig. 7 and 1st paragraph under heading "Temporal Model" re: "reference links [structural information]")

However, Auffret does not explicitly disclose:

> *scanning a hierarchical memory representation of said instance from parent description elements to child description elements until reaching the description element to be encoded, and retrieving the identification information of each scanned description element,*

North, though, discloses:

> *scanning a hierarchical memory representation of said instance from parent description elements to child description elements until reaching the description element to be encoded, and retrieving the identification information of each scanned description element,* (p. 300, Figures 14.2 and 14.3 and description between and below those figures)

It would have been obvious to one of ordinary skill in the art at the time of the invention to apply the teachings of North for the benefit of Auffret, because to do so would allow a programmer to traverse an XML document in a hierarchical fashion as taught by North in the 1st sentence under "OUTPUT Listing 14.7" on page 299. These references were all applicable to the same field of endeavor, i.e., hierarchical processing of documents.

Regarding independent method claim 3, Auffret discloses:

A decoding method for decoding a fragment comprising a content and a sequence of identification information, characterized in that it consists in:

using at least one table derived from an XML-like schema (Fig. 11), *said schema defining a hierarchical structure of description elements comprising hierarchical levels, parent description elements and child description elements* (p. 174, Fig. 7 and 1st paragraph under heading "Temporal Model" re: "graph containing description objects"), *said table containing identification information for solely identifying each description element in a hierarchical level* (p. 174, Fig. 7 and 1st paragraph under heading "Temporal Model" re: "graph containing description objects"), *and structural information for retrieving any child description element from its parent description element* (p. 174, Fig. 7 and 1st paragraph under heading "Temporal Model" re: "reference links [structural information]"),

at each step searching in said table for the description element associated to the current identification information (p. 174, Fig. 9 and subsequent description under heading "Temporal Model" re: "reference links [structural information]") *and adding said description element to a hierarchical memory representation of an instance of said schema if not already contained in said hierarchical memory representation* (p. 173, Fig. 4, and p. 175 last paragraph before the italicized heading "A *segment*"),

adding said content to the description element of said hierarchical memory representation that is associated to the last identification information of said sequence (p. 175, last paragraph before the section entitled "Relating Descriptors to an Ontology", re: tree building).

However, Auffret does not explicitly disclose:

> *scanning said sequence identification information by identification information,*

North, though, discloses:

> *scanning said sequence identification information by identification information,* (p. 300, Figures 14.2 and 14.3 and the description between and below those figures)

It would have been obvious to one of ordinary skill in the art at the time of the invention to apply the teachings of North for the benefit of Auffret, because to do so would allow a programmer to traverse an XML document in a hierarchical fashion as taught by North in the 1st sentence under "OUTPUT Listing 14.7" on page 299. These references were all applicable to the same field of endeavor, i.e., hierarchical processing of documents.

Regarding independent claim 5, Auffret discloses:

> *A encoder for encoding a description element of an instance of an XML-like schema defining a hierarchical structure of description elements* (p. 175 "Structure encoding using XML"), *said hierarchical structure comprising hierarchical levels, parent description elements and child description elements* (p. 174 Fig. 7, and paragraph under "Temporal Model" re: "graph containing description objects"), *said description element to be encoded comprising a content* (p. 175 "Structure encoding using XML"), *characterized in that it comprises:*
>
> *a memory for storing at least one table derived from said schema, said table containing identification information for solely identifying each description element in a hierarchical level* (p. 174 Fig. 7, and paragraph under "Temporal

Model" re: "graph containing description objects"), *and structural information for retrieving any child description element from its parent description element* (p. 174 Fig. 7, and paragraph under "Temporal Model" re: "reference links [i.e., structural information]"),

computing means ...

and for encoding said description element to be encoded as a fragment comprising said content and a sequence of the retrieved identification information (p. 174 "A *segment*" section, which also references Fig. 4, showing how sequenced segments are used in the building of a document).

However, Auffret does not explicitly disclose:

> *computing means for scanning said instance from parent description elements to child description elements until reaching the description element to be encoded, and retrieving the identification information of each scanned description element,*

North, though, discloses:

> *computing means for scanning said instance from parent description elements to child description elements until reaching the description element to be encoded, and retrieving the identification information of each scanned description element,* (p. 300, Figures 14.2 and 14.3 and the description between and below those figures)

It would have been obvious to one of ordinary skill in the art at the time of the invention to apply the teachings of North for the benefit of Auffret, because to do so would allow a programmer to traverse an XML document in a hierarchical fashion as taught by North in the 1st sentence under "OUTPUT Listing 14.7" on page 299. These references were all applicable to the same field of endeavor, i.e., hierarchical processing of documents.

Regarding independent claim 6, Auffret discloses:

A decoder for decoding a fragment comprising a content and a sequence of identification information, characterized in that it comprises:
a memory for storing at least one table derived from an XML-like schema (Fig. 11), *said schema defining a hierarchical structure of description elements comprising hierarchical levels, parent description elements and child description elements, said table containing identification information for solely identifying each description element in a hierarchical level* (p. 174 Fig. 7, and paragraph under "Temporal Model" re: "graph containing description objects"), *and structural information for retrieving any child description element from its parent description element* (p. 174 Fig. 7, and paragraph under "Temporal Model" re: "reference links [structural information]"),
computing means for:
... , at each step searching in said table for the description element associated to the current identification information (p. 174 Fig. 9, and subsequent description under "A *segment*") *and adding said description element to a hierarchical memory representation of an instance of said schema if not already contained in said hierarchical memory representation* (Fig. 4),
adding said content to the description element of said hierarchical memory representation that is associated to the last identification information of said sequence (p. 173 Fig. 4, and p. 175 last paragraph before section entitled "Relating Descriptors to an Ontology" re: tree building).

However, Auffret does not explicitly disclose:

> *scanning said sequence identification information by identification information,*

North, though, discloses:

> *scanning said sequence identification information by identification information,* (p. 300, Figures 14.2 and 14.3 and the description between and below those figures)

It would have been obvious to one of ordinary skill in the art at the time of the invention to apply the teachings of North for the benefit of Auffret, because to do so would allow a programmer to traverse an XML document in a hierarchical fashion as

taught by North in the 1st sentence under "OUTPUT Listing 14.7" on page 299. These references were all applicable to the same field of endeavor, i.e., hierarchical processing of documents.

Regarding independent system claim 7:

A transmission system comprising an encoder as claimed in claim 5.

Claim 7 is substantially similar to claim 5, and therefore likewise rejected.

Regarding independent system claim 8:

A transmission system comprising an decoder as claimed in claim 6.

Claim 8 is substantially similar to claim 6, and therefore likewise rejected.

20. **Claims 2 and 4 are rejected under 35 U.S.C. 103(a)** as being unpatentable over Gwendal Auffret, et al., (paper entitled: "Audiovisual-based Hypermedia Authoring: Using Structured Representations for Efficient Access to AV Documents", Hypertext '99, Darmstadt, Germany, Feb. 1999, hereafter referred to as "Auffret") in view of Simon North, et al., (SAMS Teach Yourself XML in 21 Days, Sam's Publishing, Indianapolis, IN, (c) 1999, hereafter referred to as "North") and further in view of Michael J. Hu, et al.,

(paper entitled: "Multimedia description Framework (MDF) for Content Description of Audio/Video Documents", downloaded from: arxiv.org/pdf/cs.DL/9902016.pdf, dated: Jun. 2, 1999, hereafter referred to as "Hu").

Regarding claim 2, which is dependent upon claim 1, the limitations of claim 1 have been previously addressed.

Auffret does not explicitly disclose:

> *characterized in that when a description element is defined in the schema as possibly having multiple occurrences, said table further comprises for said description element an occurrence information for indicating that said description element may have multiple occurrences in an instance, and when an occurrence having a given rank is scanned during the encoding, the corresponding retrieved identification information is indexed with said rank.*

Hu, though, discloses:

> *characterized in that when a description element is defined in the schema as possibly having multiple occurrences, said table further comprises for said description element an occurrence information for indicating that said description element may have multiple occurrences in an instance* (page 11, section 3.5, second paragraph: "Figures [sic] 7 show a .... a list [i.e., multiple occurrences] of multimedia documents ... in the content description"), *and when an occurrence having a given rank is scanned during the encoding, the corresponding retrieved identification information is indexed with said rank.* (page 11, section 3.4, second paragraph: "The target of indexing module is to automatically formulate indices of key descriptors")

It would have been obvious to one of ordinary skill in the art at the time of the invention to apply the teachings of Hu for the benefit of Auffret in view of North, because to do so would allow a user to efficiently retrieve multimedia data or documents as

taught by Hu in the first paragraph of page 11, section 3.5. These references were all applicable to the same field of endeavor, i.e., hierarchical processing of documents.

Regarding claim 4, which is dependent upon claim 3, the limitations of claim 1 have been previously addressed.

Auffret does not explicitly disclose:

> *characterized in that when a description element is defined in the schema as possibly having multiple occurrences, said table further comprises for said description element an occurrence information for indicating that said description element may have multiple occurrences in an instance, and when said sequence comprises an indexed identification information, said index is interpreted as an occurrence rank for the associated description element, same description element(s) of lower rank(s) being added to said hierarchical memory representation if not already contained in it.*

Hu, though, discloses:

> *characterized in that when a description element is defined in the schema as possibly having multiple occurrences, said table further comprises for said description element an occurrence information for indicating that said description element may have multiple occurrences in an instance* (page 11, section 3.5, second paragraph: "Figures [sic] 7 show a .... a list [i.e., multiple occurrences] of multimedia documents ... in the content description"), *and when said sequence comprises an indexed identification information, said index is interpreted as an occurrence rank for the associated description element, same description element(s) of lower rank(s) being added to said hierarchical memory representation if not already contained in it.* (page 11, section 3.4, second paragraph: "The target of indexing module is to automatically formulate indices of key descriptors")

It would have been obvious to one of ordinary skill in the art at the time of the invention to apply the teachings of Hu for the benefit of Auffret in view of North, because to do so would allow a user to efficiently retrieve multimedia data or documents as taught by Hu in the first paragraph of page 11, section 3.5. These references were all applicable to the same field of endeavor, i.e., hierarchical processing of documents.

Conclusion

21. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Non-patent Literature

Salembier, Philippe, et al., "MPEG-7: Multimedia content Description Interface, Introduction to MPEG-21 workshop, 20th and 21st of March 2000, Mar. 20 and 21, 2000, pp. 1-21.

Chiariglione, Leonardo (convenor), "Resolutions of 49th WG 11 Meeting", ISO/IEC JTC 1/SC 29/WG 11 Coding of Moving Pictures and Audio, Doc. No. N2903, downloaded from: www.itscj.ipsj.or.jp/sc29/open/29view/29n32651.pdf, dated: Oct. 1999, pp. 1-15.

Martinez, José M. (editor), "Introduction to MPEG-7 (version 1.0)", ISO/IEC JTC 1/SC 29/WG 11 Coding of Moving Pictures and Audio, Doc. No. N3545, Jul. 2000, pp. 1-10 (plus cover).

"MPEG-7: Context and Objectives (version - 10 Atlantic City)", ISO/IEC JTC1/SC29/WG 11 Coding of Moving Pictures and Audio, Doc. No. N2460, downloaded from: www.tnt.uni-hannover.de/project/mpeg/audio/public/mpeg7/w2460.pdf, dated: Oct. 1998, pp. 1-11.

Hunter, Jane, "MPEG-7 Behind the Scenes", D-Lib Magazine, vol. 5 no. 9, ISSN: 1082-9873, downloaded from: www.dlib.org/dlib/september99/hunter/09hunter.html, dated: Sep. 1999, pp. 1-12.

Vass, Jozsef, "MPEG-7: Multimedia Content Description Interface", downloaded from: meru.cecs.missouri.edu/mm_seminar/mpeg7.html, dated: Feb. 2, 1998, pp. 1-5.

Mulder, P., "The Integration of Metadata from Production to Consumer", EBU Technical Review, Sep. 2000, pp. 1-5.

Hu, Michael J., et al., "Multimedia Description Framework (MDF)for Content Description of Audio/Video Documents", DL '99, ACM 1999 1-58113-145-3/99/08, Aug. 1999, pp. 67-75 (plus citation sheet).

Staab, Steffan, "Intelligent Systems on the World Wide Web: 11 MPEG-7 Lecture Slides", downloaded from: "www.aifb.uni-karlsruhe.de/WBS/sst/Teaching/Intelligente System im WWW SS 2000/11-MPEG.pdf", dated: Aug. 23, 2000, slides 1-15 (plus Wayback Internet archive page and screen capture of pdf file URL).

US Patent Application Publications

Jain et al	US2001/0018693

US Patents

Mohan et al	6,748,382
Srivastava et al	6,549,922
Jain et al	6,360,234
Huang et al	6,593,936
Bergman et al	6,564,263
Sezan et al	6,236,395
DaGraca et al	6,646,676
Kumar et al	6,665,731

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Robert M Stevens whose telephone number is (703) 605-4367. The examiner can normally be reached on M-F 7:00 - 3:30. After mid-October 2004, the Examiner can be reached at (571) 272-4102.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Joseph Feild can currently be reached on (703) 305-9792. The current fax phone number for the organization where this application or proceeding is assigned is 703-872-9306. However, note that the main number for Technology Center 2100 will be (571) 272-2100, as of mid-October 2004.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Robert M. Stevens
Art Unit 2176
Date: September 22, 2004

JOSEPH FEILD
SUPERVISORY PATENT EXAMINER

rms

Sheet 1 of 1

10000 U.S. PTO 09/982269 10/17/01

Form PTO-1449 U.S. DEPARTMENT OF COMMERCE (REV.7-80) PATENT AND TRADEMARK OFFICE	Atty. Docket No. FR 000110	Serial No.
	Applicant BENOIT MORY ET AL	
INFORMATION DISCLOSURE CITATION (Use several sheets if necessary)	Filing Date CONCURRENTLY	Group

U.S. PATENT DOCUMENTS

Ex. Int.		Document Number	Date	Name	Class	Sub-class	Filing Date If Approp.
	AA						
	AB						
	AC						
	AD						
	AE						
	AF						

FOREIGN PATENT DOCUMENTS

		Document Number	Date	Country	Class	Sub-class	Trans. Yes	Trans. No
	AG							
	AH							
	AI							
	AJ							
	AK							

OTHER (Including Author, Title, Date, Pertinent Pages, Etc.)

RS	AL		"Everything You Wanted to Know About MPEG=7: Part 2, Frank Nack, GMD-IPSI, Adam T. Lindsay, (Starlab) 1999, IEEE, pp. 64-73.
	AM		
	AN		

Examiner Robert Stevens Date Considered 9/17/04

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; Draw lir through citation if not in conformance and not considered. Include a copy of this form with next communication to applicant.

Form PTO-1449 (REV. 7-80) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	Atty. Docket No. PHFR 000110	Serial No. 09/982,269
	Applicant BENOIT MORY ET AL	
	Filing Date OCTOBER 17, 2001	Group 2176

INFORMATION DISCLOSURE CITATION
(Use several sheets if necessary)

U.S. PATENT DOCUMENTS

Ex. Int		Document Number	Date	Name	Class	Sub-class	Filing Date If Approp.
	AA						
	AB						
	AC						
	AD						
	AE						
	AF						

FOREIGN PATENT DOCUMENTS

		Document Number	Date	Country	Class	Sub-class	Trans. Yes	Trans. No
KS	AG	WO 97/42040	03/1997	PCT (WORLD)	G06F	15/00		X
	AH							
	AI							
	AJ							
	AK							

OTHER (Including Author, Title, Date, Pertinent Pages, Etc.)

	AL		
	AM		
	AN		

Examiner Robert Stevens Date Considered 9/17/04

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with M 609; Draw line through citation if not in conformance and not considere Include a copy of this form with next communication to applicant.

Notice of References Cited	Application/Control No. 09/982,269	Applicant(s)/Patent Under Reexamination MORY ET AL.	
	Examiner Robert M Stevens	Art Unit 2176	Page 1 of 3

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
	A	US-6,748,382	06-2004	Mohan et al.	707/10
	B	US-6,549,922	04-2003	Srivastava et al.	707/205
	C	US-6,360,234	03-2002	Jain et al.	715/500.1
	D	US-6,593,936	07-2003	Huang et al.	345/619
	E	US-6,564,263	05-2003	Bergman et al.	709/231
	F	US-6,236,395	05-2001	Sezan et al.	345/723
	G	US-2001/0018693	08-2001	JAIN et al.	707/500
	H	US-6,646,676	11-2003	DaGraca et al.	348/155
	I	US-6,665,731	12-2003	Kumar et al.	709/246
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	Salembier, Philippe, et al., "MPEG-7: Multimedia content Description Interface, Introduction to MPEG-21 workshop, 20th and 21st of March 2000, Mar. 20 and 21, 2000, pp. 1-21.
	V	Chiariglione, Leonardo (convenor), "Resolutions of 49th WG 11 Meeting", ISO/IEC JTC 1/SC 29/WG 11 Coding of Moving Pictures and Audio, Doc. No. N2903, downloaded from: www.itscj.ipsj.or.jp/sc29/open/29view/29n32651.pdf, dated: Oct. 1999, pp. 1-15.
	W	Martinez, José M. (editor), "Introduction to MPEG-7 (version 1.0)", ISO/IEC JTC 1/SC 29/WG 11 Coding of Moving Pictures and Audio, Doc. No. N3545, Jul. 2000, pp. 1-10 (plus cover).
	X	"MPEG-7: Context and Objectives (version - 10 Atlantic City)", ISO/IEC JTC1/SC29/WG 11 Coding of Moving Pictures and Audio, Doc. No. N2460, downloaded from: www.tnt.uni-hannover.de/project/ mpeg/audio/public/mpeg7/w2460.pdf, dated: Oct. 1998, pp. 1-11.

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

Notice of References Cited	Application/Control No. 09/982,269	Applicant(s)/Patent Under Reexamination MORY ET AL.	
	Examiner Robert M Stevens	Art Unit 2176	Page 2 of 3

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
	A	US-			
	B	US-			
	C	US-			
	D	US-			
	E	US-			
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	Hunter, Jane, "MPEG-7 Behind the Scenes", D-Lib Magazine, vol. 5 no. 9, ISSN: 1082-9873, downloaded from: www.dlib.org/dlib/september99/hunter/09hunter.html, dated: Sep. 1999, pp. 1-12.
	V	Vass, Jozsef, "MPEG-7: Multimedia Content Description Interface", downloaded from: meru.cecs.missouri.edu/mm_seminar/mpeg7.html, dated: Feb. 2, 1998, pp. 1-5.
	W	Mulder, P., "The Integration of Metadata from Production to Consumer", EBU Technical Review, Sep. 2000, pp. 1-5.
	X	Auffret, Gwendal, et al., "Audiovisual-based Hypermedia Authoring: Using Structured Representations for Efficient Acess to AV Documents", Hypertext '99, ACM 1999 1-58113-064-3/99/2, Feb. 1999, pp. 169-178 (plus citation sheet).

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

Notice of References Cited	Application/Control No. 09/982,269	Applicant(s)/Patent Under Reexamination MORY ET AL.	
	Examiner Robert M Stevens	Art Unit 2176	Page 3 of 3

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
	A	US-			
	B	US-			
	C	US-			
	D	US-			
	E	US-			
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	Hu, Michael J., et al., "Multimedia Description Framework (MDF)for Content Description of Audio/Video Documents", DL '99, ACM 1999 1-58113-145-3/99/08, Aug. 1999, pp. 67-75 (plus citation sheet).
	V	Hu, Michael J., et al., "Multimedia Description Framework (MDF)for Content Description of Audio/Video Documents", downloaded from: arxiv.org/pdf/cs.DL/9902016.pdf, dated: Jun. 2, 1999, pp. 1-20.
	W	Staab, Steffan, "Intelligent Systems on the World Wide Web: 11 MPEG-7 Lecture Slides", downloaded from: "www.aifb.uni-karlsruhe.de/WBS/sst/Teaching/Intelligente System im WWW SS 2000/11-MPEG.pdf", dated: Aug. 23, 2000, slides 1-15 (plus 2 date pages).
	X	North, Simon, et al., SAMS Teach yourself XML in 21 Days, Sam's Publishing, Indianapolis, IN, © 1999, pp. 291-305, 343-359 and 361-373.

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

Introduction to MPEG-21 workshop
20th & 21st of March 2000

MPEG-7: Multimedia Content Description interface

Philippe Salembier
Universitat Politècnica de Catalunya
philippe@gps.tsc.upc.es

Olivier Avaro
France Telecom
olivier.avaro@cnet.francetelecom.fr

Outline

- Goals and scope of MPEG-7
- MPEG-7 Tools and Architecture
- MPEG-7 and MPEG-21
- Conclusions

Objective of MPEG-7

- Standardize content-based description for various types of audiovisual information
  - allow quick and efficient content localization
  - address a wide range of applications

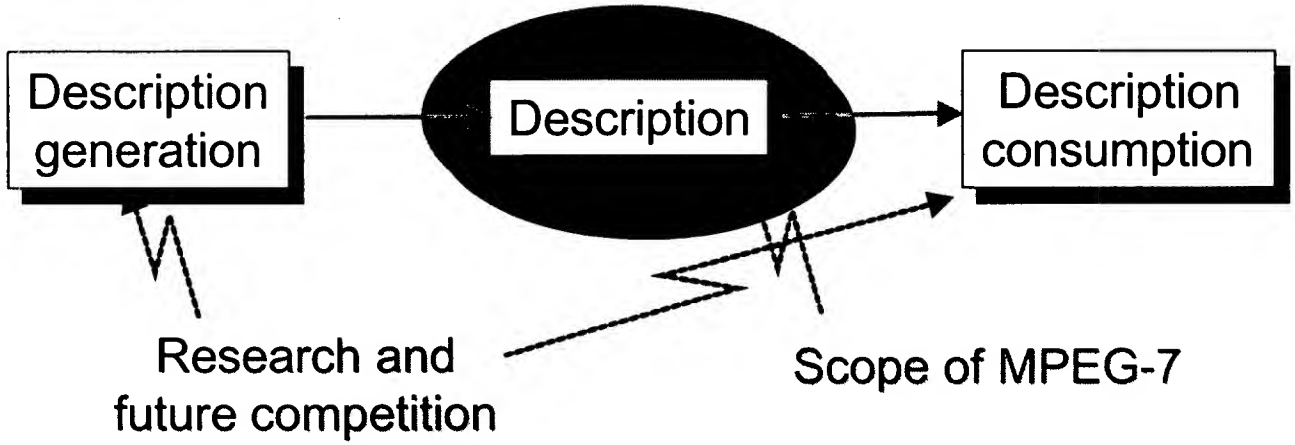
  ⇒ For some applications: need to go beyond the pure content description, e.g, user preferences

- Types of audiovisual information:
  - Audio, speech
  - Moving video, still pictures, graphics, 3D models
  - Information on how objects are combined in scenes
- Descriptions independent of the media support
- Existing solutions for textual content or description

Example of application areas

- Storage and retrieval of audiovisual databases
- Broadcast media selection (Radio, TV programs)
- Personalized news service on Internet (push media filtering)
- Tele-shopping
- Intelligent multimedia presentations
- Educational applications
- Surveillance and remote sensing
- Bio-medical applications
- etc.

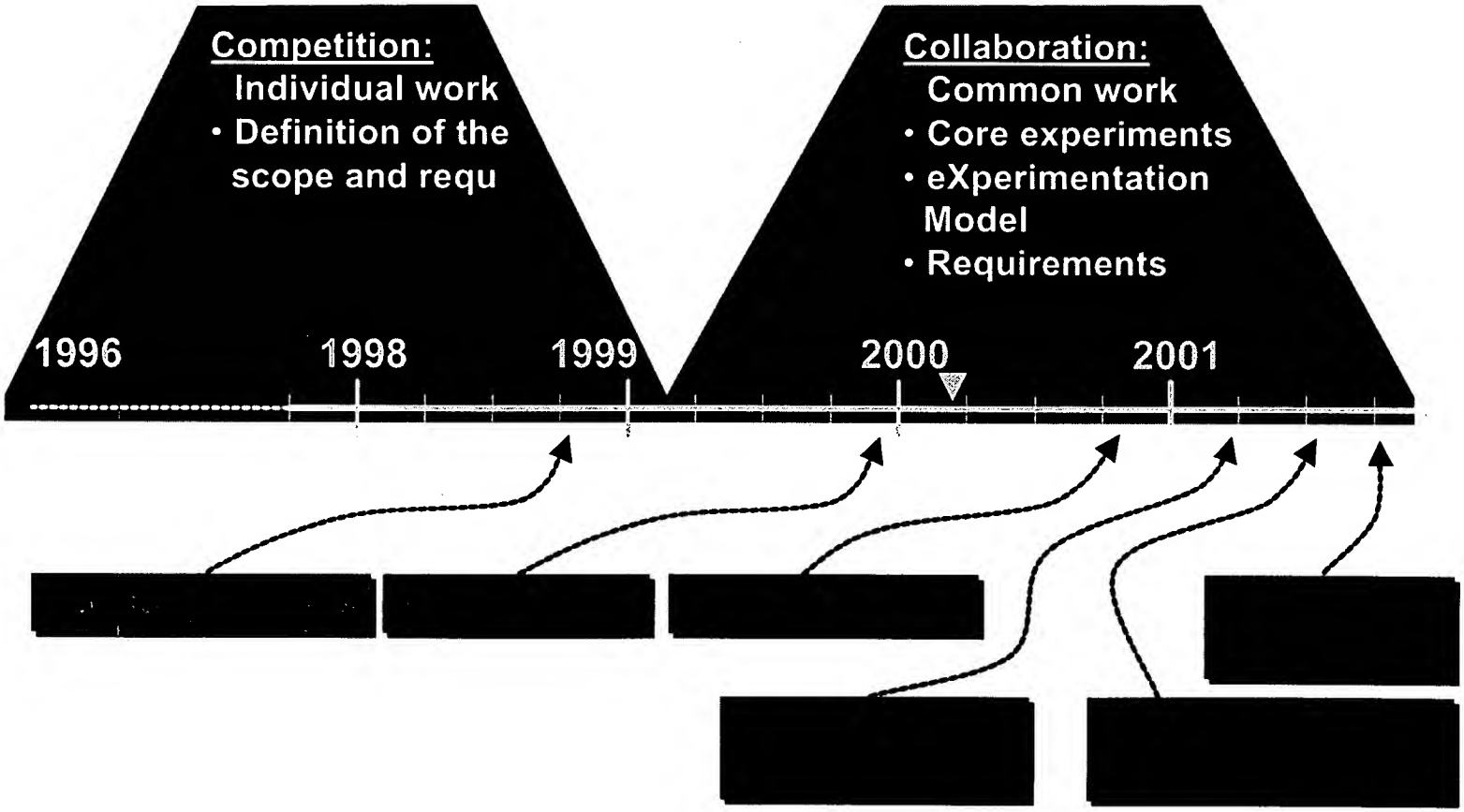
Scope of MPEG-7



- The description generation (feature extraction, indexing process, etc.) and consumption (search engine, retrieval process, etc.) are not a normative part of MPEG-7

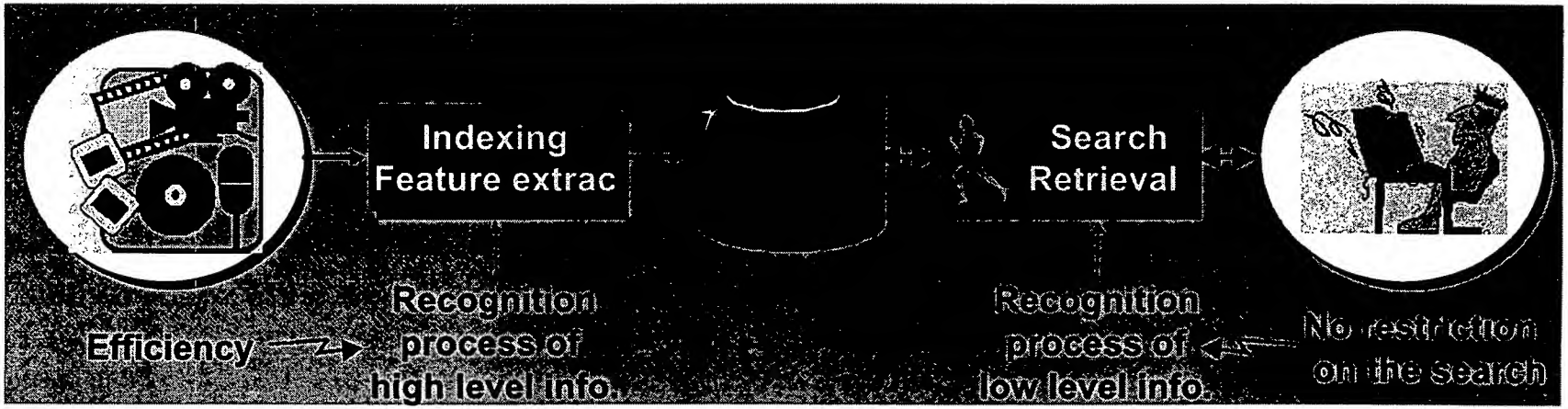
⇒ Just the description is normative

Workplan


Competition:
Individual work
• Definition of the scope and requ
Collaboration:
Common work
• Core experiments
• eXperimentation Model
• Requirements
1996
1998
1999
2000
2001

Description

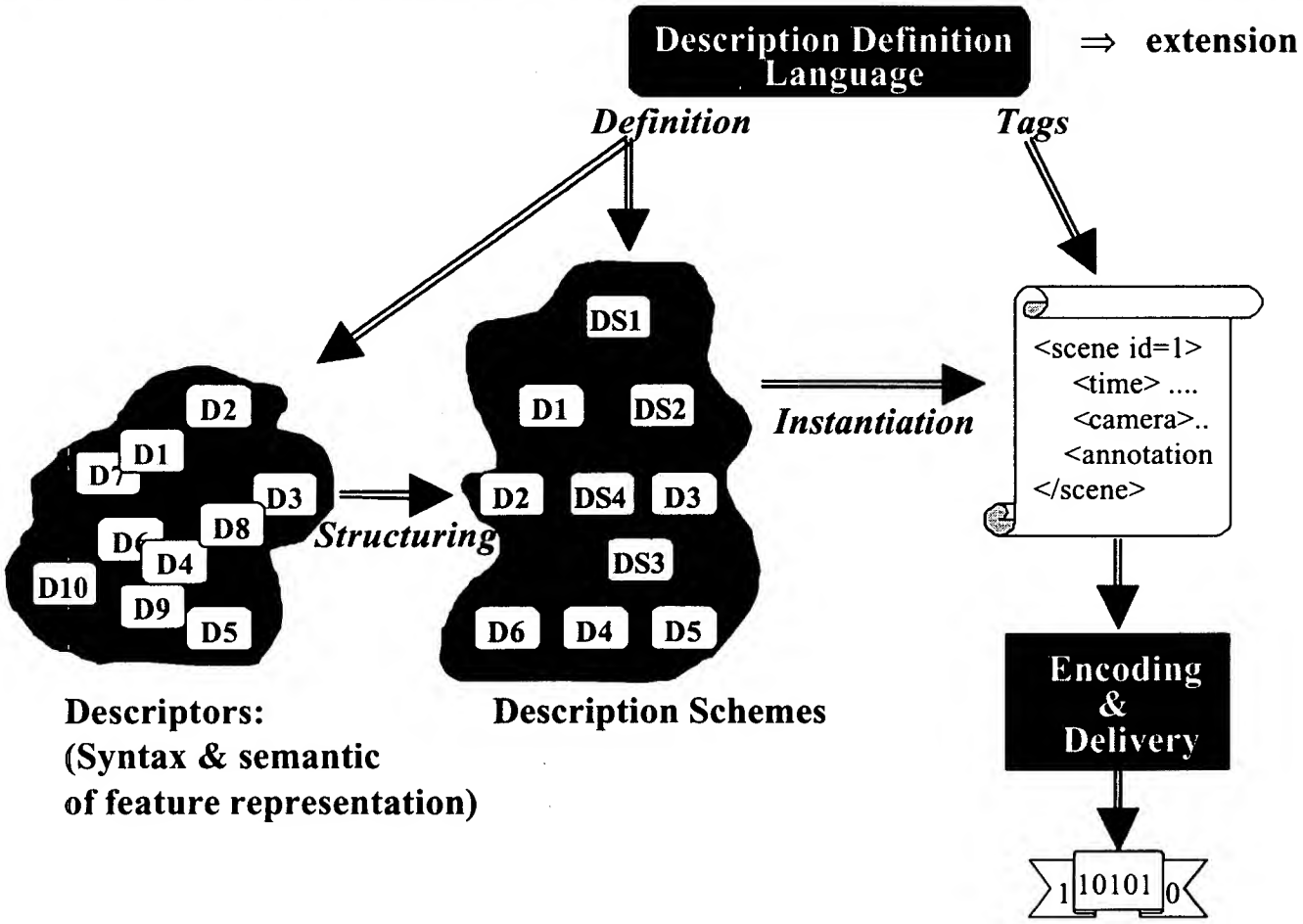
- Information about the content: recording date & conditions, title, author, copyright, coding format, classification, etc.
- Information present in the content: Combination of low level and high level descriptors
  - High level description:
    - Efficient and powerful
    - Lack of flexibility
  - Low level description
    - Generic and flexible
    - Intelligent / efficient search engine



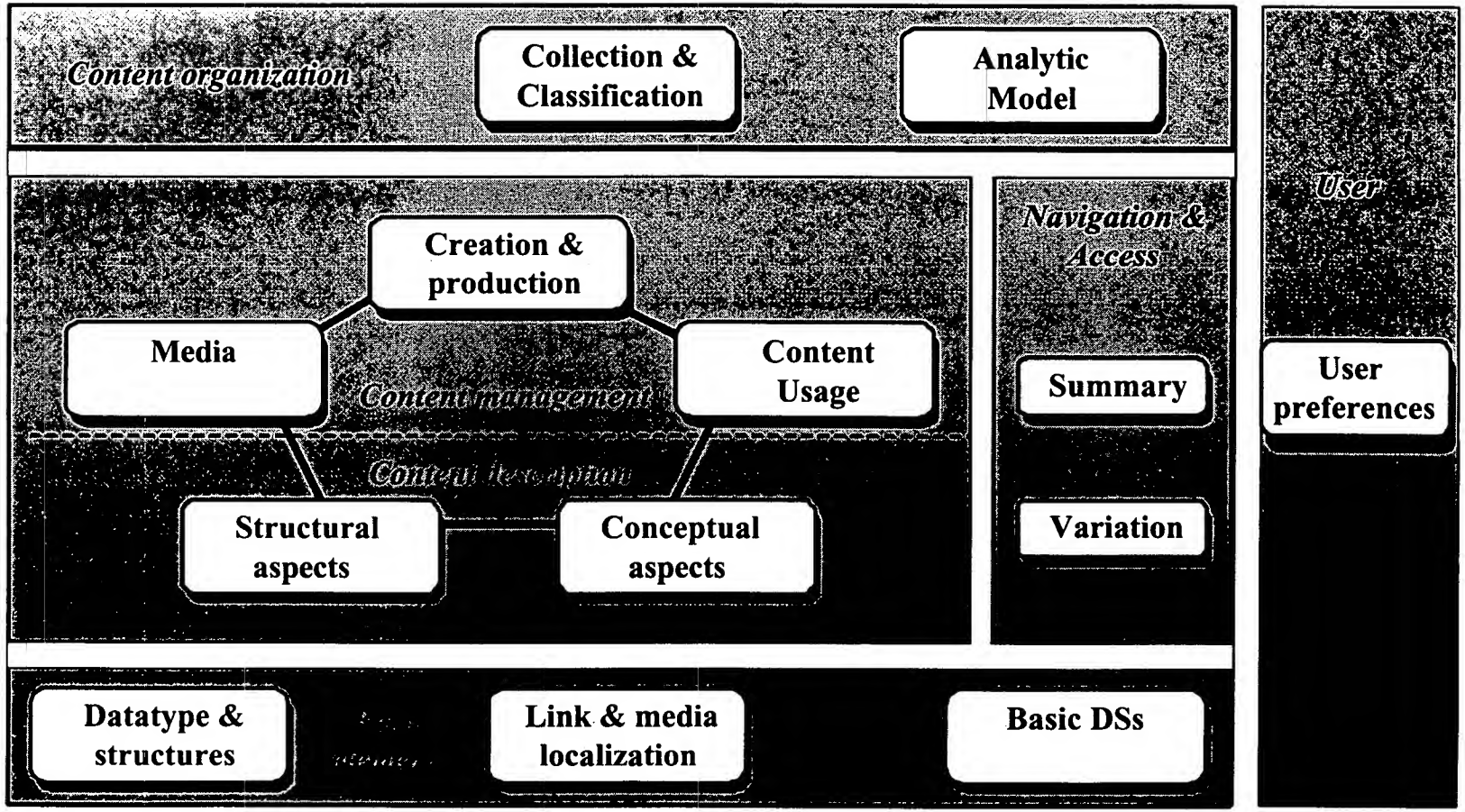
Outline

- Goals and scope of MPEG-7
- MPEG-7 Tools and Architecture
- MPEG-7 and MPEG-21
- Conclusions

Current MPEG-7 working areas



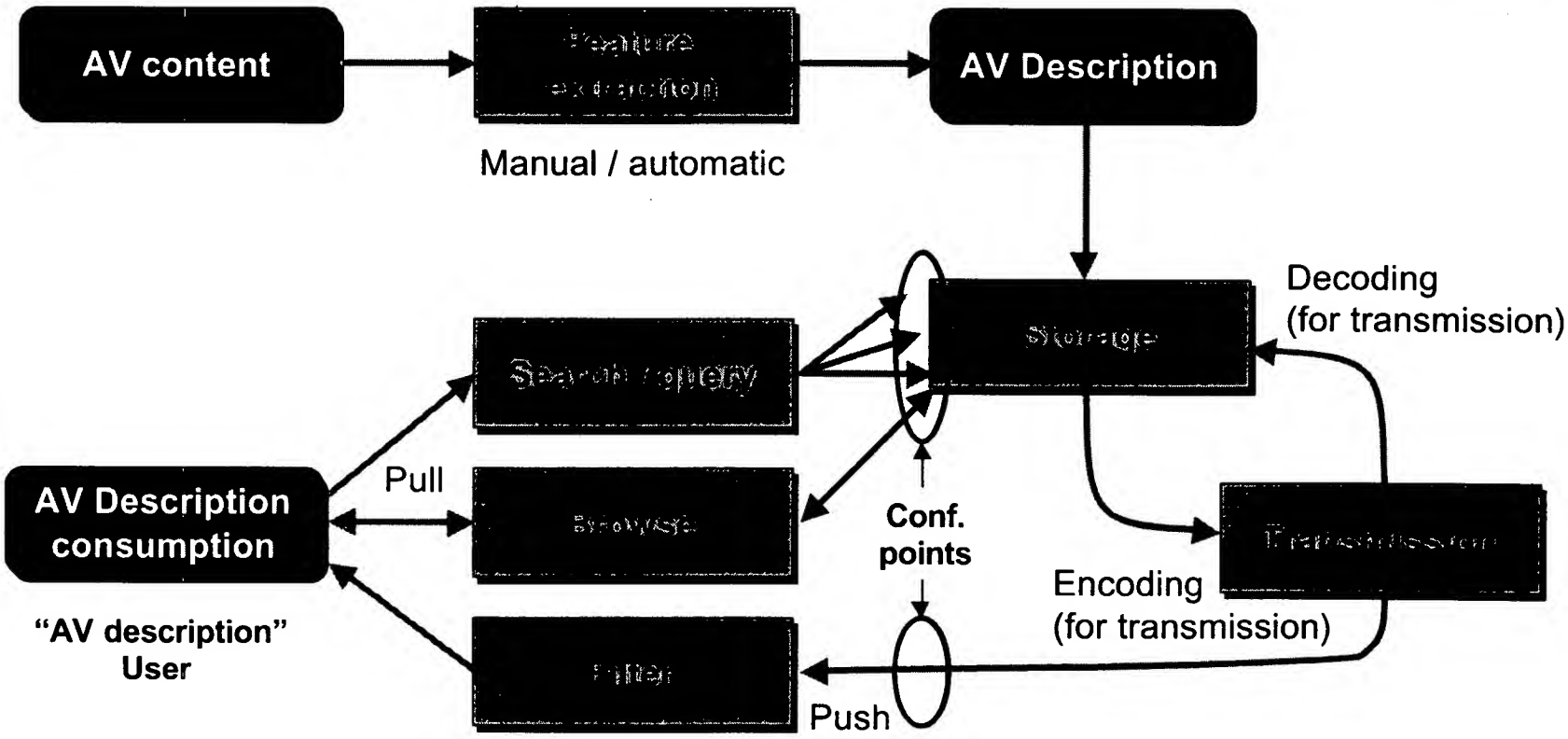
Library of description tools



Use of description tools

- The description tools are presented on the basis of the functionality they provide.
- In practice, they are combined into meaningful sets of description units.
- Furthermore, each application will have to select a sub-set of descriptors and DSs.
- Library of tools!
- DDL can be used to handle specific needs of the application.

An example of architecture



- The content and its description may also be multiplexed

Outline

- Goals and scope of MPEG-7
- MPEG-7 Tools and Architecture
- MPEG-7 and MPEG-21
- Conclusions

Key-areas of the Multimedia framework

From a user viewpoint (N3162):

- Network and delivery
- Quality and flexibility of service **MPEG-7**
- Quality of content (rendering)
- Quality of content (artistic)
- Ease of use of services and devices
- Interoperability of physical media formats
- Payment / subscription model

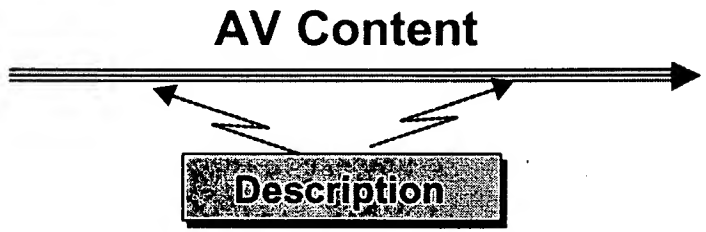
Multi-platform decoding and rendering

Search, filter, locate, retrieve and store the content MPEG-7

- Consumer content publishing **MPEG-7**
- Usage rights of consumers
- Consumer privacy

Content representation / description

- Generally seen as two different entities
  - Two different functionalities:
  - Different set of requirements
  - Different tools, etc.



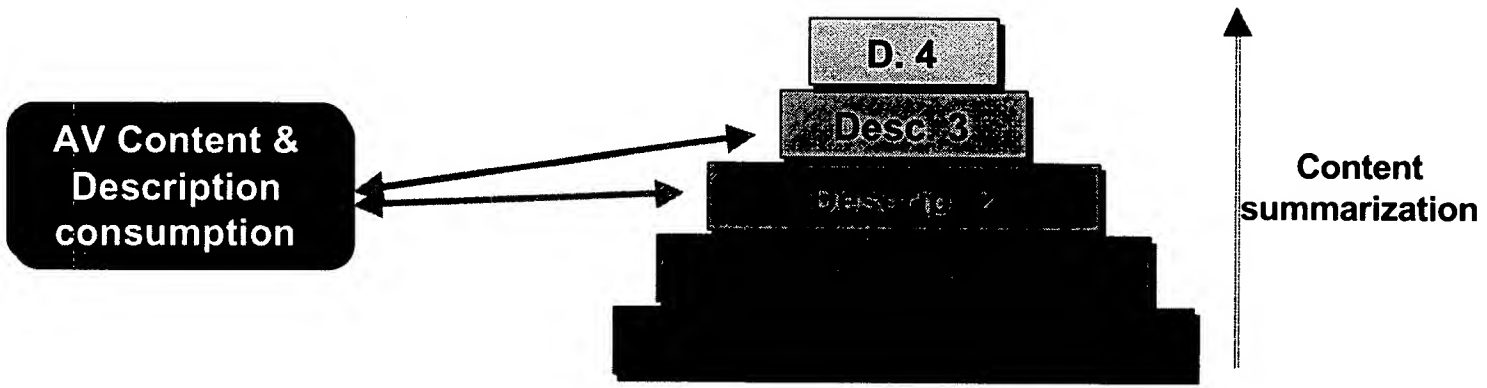
Content representation / description

- For future multimedia services, content representation and description may have to be addressed jointly:
  - Many services dealing with content representation will have to deal first with content description

    "a non-described content may be useless"
  - Access only to the content description:
    - New original services (e.g. optimizing personal time)
    - Adaptation to networks and terminal capabilites

Content representation / description

- Is consumption of the content the ultimate goal ?



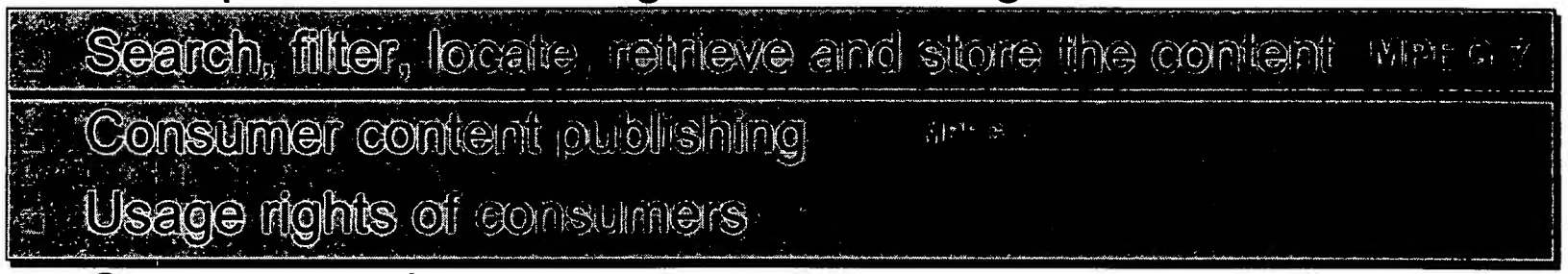
Integration of content representation / description

- Full integration of content representation and its description:
  - Are there new requirements?
  - What is the appropriate architecture for this integration?
- Analysis of the integration between content representation and its description: a work item for MPEG-21?

Key-areas of the Multimedia framework

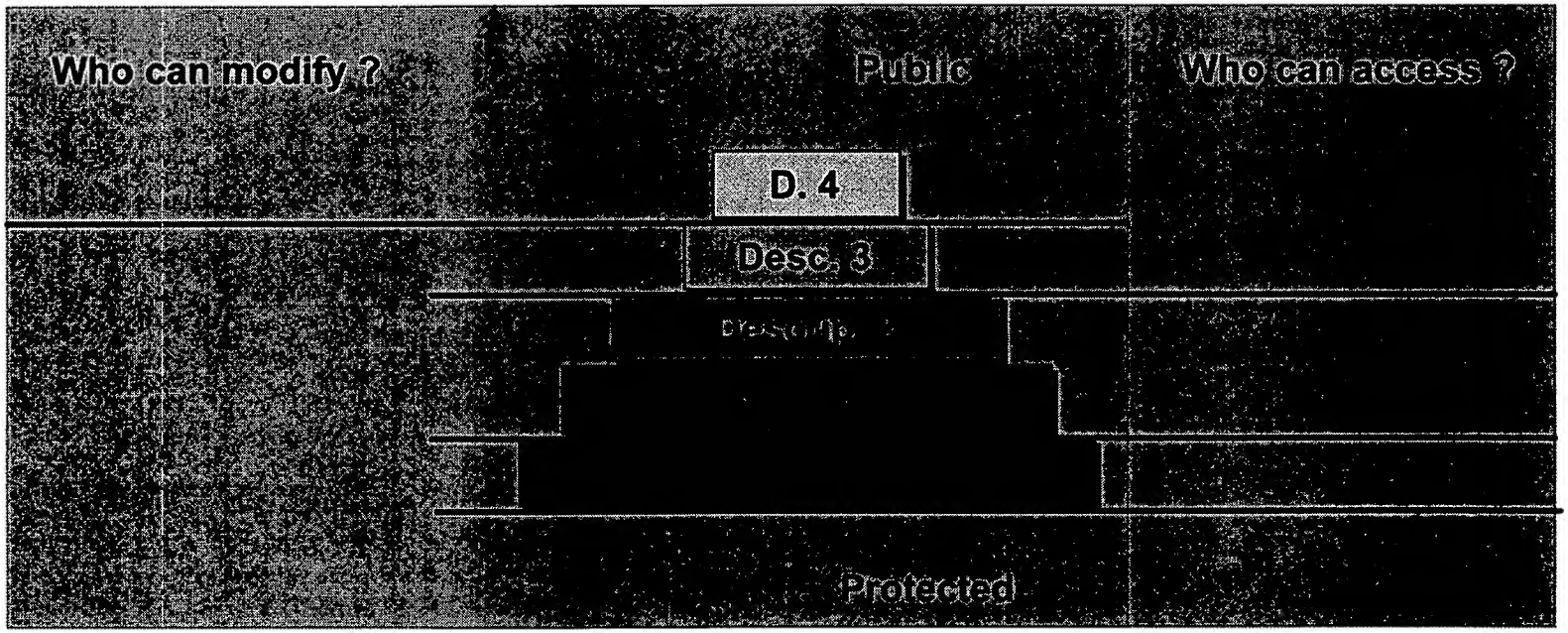
From a user viewpoint (N3162):

- Network and delivery
- Quality and flexibility of service **MPEG-7**
- Quality of content (rendering)
- Quality of content (artistic)
- Ease of use of services and devices
- Interoperability of physical media formats
- Payment / subscription model
- Multi-platform decoding and rendering



- Consumer privacy

Information Protection



- The definition of the frontier between public/protected information will create / allow different business models.

Conclusions

- Overview of MPEG-7
  - Scope and goals, Tools, Architecture
- Relation: content representation and content description
  - No clear separation for future services
  - Integration of content representation and description may be studied by MPEG-21
- Relation: content description and protection
  - The definition of what is protected will enable different business models
  - The value of the description is going to increase

INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/SC 29/WG 11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC 1/SC 29/WG 11 **N2903**

October 1999

Source: Leonardo Chiariglione – Convenor
Title: Resolutions of 49th WG 11 meeting
Status:

1 WG11 approves the reports from the Requirements, Delivery, Systems, MDS, Video, Audio, SNHC, ISG, Liaison and HoD groups.

2 Subgroup recommendations

2.1 The Requirements group

2.1.1 recommends to approve the following documents:

Name	*No.*
MPEG-4 Requirements Document, V.12	2992
MPEG-4 New Profiles under Consideration	2993
Study on V.2 MPEG-4 Audio FPDAM	2946
MPEG-4 Overview	2995
MPEG-7 Requirements Document V.10	2996
MPEG-7 DDL development document V.2	2997
Overview of MPEG-7 Descriptors and Description Schemes V.0.2	2998
MPEG-7 Development Process	2999
MPEG-7 Intellectual Property Management and Protection, V.0.1	3001
First ideas on Multimedia Framework	3002

2.1.2 recommends to establish the following AdHoc groups :

Title		*No.*	*Mtg*
AHG for Study on potential new MPEG-2 Levels and Progressive Profile(s)	Craig Birkmaier	3003	yes
AHG on DDL development	J. Hunter	3004	yes
AHG on IPMP in MPEG-4	N. Rump, H. Inoue, Y. K. Chang	3034	no
AHG on MPEG-7 Intellectual Property Management & Protection	N. Rump, K. Hill	3006	no
AHG on Multimedia Framework	K. Hill, R. Koenen	3007	no

2.1.3 recommends to make publicly available the following documents :

Title	*No.*
MPEG-4 Requirements Document V.12	2992
MPEG-4 Overview	2995
MPEG-7 DDL development document V.2	2997
MPEG-7 Requirements Document V.10	2996

2.1.4 The Requirements Group thanks companies and national bodies that have submitted Profile support statements.

2.1.5 The Requirements Group intends to make statements of support for Profiles and their intended deployment public on the MPEG homepage after the Maui meeting. This will happen in the form of making public the Profiles under Consideration document, or an excerpt of that document. Companies that rather wish not to be mentioned publicly are kindly requested to inform MPEG and to withdraw their support in a written input to the 50th MPEG meeting.

2.1.6 The Requirements group understands that companies that support a profile to be included into the MPEG-4 Standard, commit to doing the conformance work for the profile. This includes generating test bitstreams for each of the desired Levels. Companies that cannot commit to doing conformance for a profile are kindly requested to withdraw their support for such a profile.

2.1.7 The Requirements Group asks companies supporting scalable visual profiles for Version 2 to reconsider their requests in the light of the work on Fine Granular Scalability, and to see whether their needs can be met by FGS, or by making a combination of FGS and other types of scalability. The reason for this request is the desire to keep the amount of profiles limited.

2.1.8 The Requirements Group informs National Bodies that suggestions for new Profiles and Levels can be found in the document 'MPEG-4 Profiles Under Consideration', N2993, and asks NBs to use this document as a reference when they cast their votes on the Version 2 FPDAM ballots. Note that some of the Level definitions in this document also apply to Profiles already in the standard (so-called 'V.1 Profiles'), and that these elements are added in the amendment known as Version 2.

2.1.9 Having heard contributions on the Studio Profile(s), the Requirements group agrees with the desirability to have 1) normative definition of the I-DCT and 2) the possibility of having an 'uncoded' block mode.

2.1.10 The Requirements Group confirms the desirability of being able to combine low latency sprites with grey scale shape in MPEG-4 Visual, and advises that the syntax be changed so as to allow this combination, as soon as possible.

2.1.11 The Requirements group recommends that MPEG reassesses it position with respect to standardising digital rights management technology, considering that new technology is emerging, and that protecting digital assets is of increasing importance for the success of the MPEG-4 standard.

2.1.12 The Requirements group recommends that the DDL work, which has clearly left the requirements stage, be continued in the context of the Systems Group as from the next meeting.

2.1.13 The Requirements group recommends that a conceptual model of MPEG-7 Descriptors and Description Schemes be developed and maintained, and that proposers of new technology attempt to address the relation of their proposal with this conceptual model in their submissions, in accordance with the guidelines found in N2999 (MPEG-7 Development Process)

2.1.14 Frank Nack is thanked for his successful efforts in leading the DDL work. Best wishes

for his future endeavours.

2.1.15 That the following output documents from previous meetings remain on the MPEG home page for public access:

MPEG-4 Applications Document	*N2563*
Results of AAC Subjective tests	*N2006*
Overview of MPEG-4 Profile and Level definitions	*N2458*
MPEG-4 Audio verification test results: speech codecs	*N2424*
MPEG-4 Audio verification test results: audio on Internet	*N2425*
MPEG-4 Video verification tests: error resilience	*N2604*
MPEG-4 Video verification tests: temporal Scalability in Simple Scalable Profile	*N2605*
MPEG-4 Video verification tests: Content Based Coding	*N2711*
MPEG-4 Intellectual Property Management & Protection (IPMP) Overview & Applications	*N2614*
Description of MPEG-7 Content Set	*N2467*
Licensing Agreement for MPEG-7 Content Set	*N2466*
Guide to obtaining the MPEG-7 Content Set	*N2570*

2.2 The Systems group recommends

2.2.1 the approval of the following documents:

Title	No.
DoC on 13818-6 FDAM 1	N2916
Text of 13818-6 FPDAM 1.2	N2917
Study of FPDAM 7 of MPEG-2 Systems	N2910
Disposition of comments on FPDAM 7 of MPEG-2 Systems	N2909
Study of Conformance FCD	N2920
Work plan for Systems conformant bitstream production	N3018
Corrigendum for 14496-1	N3019
Study of FPDAM 1 of 14496-1 (MPEG-4 Systems Version 2)	N3020
Study on Internet draft for the carriage of MPEG-4 on IP	N3021
Revised Text for WG 11 N2873 Draft Agreement with Sun Microsystems	N3022
Text of ISO/IEC 14496-5 PDAM 1	N2918
Text of ISO/IEC 14496-6 DCOR 1	N3023
Study of FPDAM 1 of 14496-6 (MPEG-4 DMIF Version 2)	N3024
MPEG-4 Systems Version 2 VM 8.0	N3025
Status of the Systems Version 1&2 Software Implementation	N3026
Systems Software Implementation Work plan	N3027
Template for NB Comments	N3028
DMIF FAQ Version 4.0	N3029

2.2.2 In the process of further verifying its reference software, the Systems sub-group has discovered errors in the code related to the script node. In case the proponents of this technology or another company interested in the technology do not help fixing the errors, the Systems sub-group recommends that this technology be removed from the 14496-1 through a corrigendum of 14496-1.

2.2.3 The Systems sub-group acknowledges the commitments of the following companies to

provide test streams for MPEG-4 Systems conformance: FT R&D, IBM, ENST, Optibase, HUT, Sun Microsystems Inc., Valley Consultants, and CSELT according to the work plan described in N3018 and encourages other companies to join this effort.

2.2.4 The Systems sub-group acknowledges the reception of document M5229 presenting the proposed guidelines *regarding normative references to external specification* documents. As the 14496-1/Amd 1 (MPEG-J) has been initiated prior to the JTC1 directive, the guidelines described in document M5229 can not be enforced to apply. Nevertheless, the Systems sub group will exert is best efforts to come as close as possible to these guidelines within the current 14496-1/Amd 1 (MPEG-J) timeline.

2.2.5 The Systems sub-group acknowledges the reception of document M5096 proposing possible improvement of N2873 (Draft Agreement with Sun Microsystems). All the comments proposed have been approved and included in N3022.

2.2.6 That National Bodies should comment on the design within MPEG of a textual language for MPEG-4 content authoring.

2.2.7 That the following companies: AT&T, CRL, CSELT, ENST, IBM, Nokia, Thomson-CSF, Samsung, UCL should be thanked for bringing compelling demonstrations of MPEG-4 applications.

2.2.8 That some IPMPS_Type are reserved for ISO specific use and that further studies are conducted within the IPMP AHG to evaluate if the IPMP descriptor extensions proposed by the MPEG-PF is needed to be included within the normative part of the MPEG-4 standard.

2.2.9 That the following documents: N3020, N3025 will be released within 14 days after the Melbourne meeting to accommodate final text editing.

2.2.10 That the following documents: N2910, N3024 will be released within 21 days after the Melbourne meeting to accommodate final text editing.

2.2.11 to establish the following AdHoc Groups:

Title	Chair(s)	No.
AHG on Systems Conformance	Dufourd & al.	N3030
AHG on Advanced BIFS	Signes & al.	N3031
AHG on MPEG-J	Swaminathan & al.	N3032
AHG on MPEG-4 File Format	Singer & al	N3033
AHG on Intellectual Property Management & Protection within MPEG-4	Rump & al.	N3034
AHG on IM 1	Lifshitz & al.	N3035
AHG on MPEG-4 Content on MPEG-2 Systems and on the Internet	Carsten Herpel, Jan Van Der Meer, S. Casner	N3036
AHG on Multi-user applications	Olivier Avaro & al.	N3037
AHG on Back Channel and ESM	Young-Kwon Lim & al.	N3038
AHG on MPEG-7 Systems	O. Avaro et al.	N3039
AHG on MPEG-7 Linking	Joerg Heuer et al.	N3040
AHG on Multiple DII messages in download protocol	Matt Goldman (DiviCom)	N3041

2.2.12 to make publicly available the following documents:

Title	No.
Study of FPDAM 7 of MPEG-2 Systems	N2910
Corrigendum for 14496-1	N3019
Study on Internet draft for the carriage of MPEG-4 on IP	N3021
Corrigendum for 14496-6	N3023
DMIF FAQ Version 4.0	N3029

2.3 The Multimedia Description Scheme group recommends:

2.3.1 recommends to approve the following documents:

Title	*No.*
MPEG-7 Generic AV Description Schemes (V 0.7)	N2966
Supporting information for the Generic AV Description Schemes	N2967
MPEG-7 DS Validation Experiment on the Syntactic DS, the Semantic DS, and the Syntactic/Semantic Link DS	N2968
Validation of the MPEG-7 Summary DS	N2969
Validation Experiment for MPEG-7 Model-DS on Visual Data	N2970
Validation Experiments for Universal Multimedia Access	N2971
MPEG-7 Core/Validation Experiment on the Weight DS	N2972
Validation Experiment on MPEG-7 DS from the Viewpoint of Video Editing and Production	N2973
Digital Patient Record Validation Experiment for MPEG-7 Description Schemes	N2974
Validation Experiment for Ordered Relation Graphs	N2975
Description of Validation Experiments for MPEG-7 SpatioTemporalRegion DS	N2928

2.3.2 recommends to establish the following AdHoc groups :

Title	Chair(s)	*No.*	*Mtg*
AHG on MPEG-7 Generic Description Schemes Development	P. Salembier, S. Quackenbush, C. Saraceno, S. Jeannin, T. Walker	N2976	Yes
AHG on Media and Meta DSs and Harmonization with other Schemes	J. Martinez, M. Cox,	N2977	Yes
AHG on DS Validation and Core Experiment	A. Benitez, N. Day, S. Devillers, Jin Soo Lee	N2978	Yes
AHG on Generic AV DS Conceptual Model	J. Smith, H. Rising, U. Srinivasan	N2979	Yes
Ah Hoc Group on Semantic DS	R. Leonardi, M. Tekalp	N2980	Yes
Ad Hoc Group on User Preferences in MPEG-7	I. Sezan, K. Yoon	N2981	No
Ad Hoc Group on MPEG-7 Linking	J. Heuer , E. Wan, O. Avaro	N3040	Yes

2.3.3 To encourage the development of description schemes describing processes in MPEG-7, in particular: Acquisition, Delivery, Editing, Presentation, Production, and Publication.

2.3.4 To encourage the development of Audio and Semantic related Description Schemes. To actively work on the harmonization of the Meta / Media Description Schemes with

relevant standards including SMPTE Metadata Dictionary, Dublin Core, INDECS.

2.3.5 To extend to current MPEG-7 content set to include in particular video content in English, video content with multiple video and/or audio tracks, video content with text captions in one or more languages, other multimedia content such as web pages (with audio, video, sound, and text) and multimedia presentations (SMIL and MPEG-4), etc. This content should be used under the conditions described in document N2466: "Licensing Agreement for the MPEG-7 Content Set".

2.4 The Video group recommends

2.4.1 to thank the members of the Japanese National Body who have generously offered to donate video conformance bitstreams, which will be evaluated by the Video Sub-group for incorporation in ISO/IEC 14496-4;

2.4.2 to thank Intel Corp for their generous offer to donate object code for the MPEG-7 XM software, but to decline this offer so as to ensure that source code for all parts of the XM software is available to MPEG members;

2.4.3 that future contributions on watermarking will only be accepted if they justify clearly to the Requirements Sub-group why the proposed technology should be the subject of standardisation by MPEG, and strongly discourages proposals that do not provide this justification;

2.4.4 to issue at the next meeting in Maui a call for coding technology to be compared to existing and emerging MPEG 4 video standards in subjective tests, in order to confirm that MPEG-4 video provides state of the art technology or to identify suitable technology for extensions;

2.4.5 that WG11 thank the USNB for its comment on the subjective test document N2824 and to inform the USNB that the Main Profile could have performed better in the tests if the sprite tool had been included. This was not done because the sprite tool is considered too complex for the target application of the ACE profile.

2.4.6 to make MPEG-2 Video elementary bitstreams publicly available to allow for interested parties to self-assess the impact of using extension start codes identifiers "1101".

2.4.7 To request the help of the ISG at the Maui meeting to evaluate the performance on software and hardware complexity of fast motion estimation algorithms

2.4.8 to approve the following documents :

Title	No.
Text of ISO/IEC 14496-2 /DCOR1	2919
Study of ISO/IEC FCD 14496-4 (Video)	2920
Study of ISO/IEC 14496-2 /FPDAM1	2921
Study of ISO/IEC 14496-5 FDIS	2922
Response to comment by AT NB on ISO/IEC 14496-2 /FDIS	2923
Description of Core Experiments in FGS	2924
Text of ISO/IEC 14496-2 FGS Amendment WD 2.0	2925
Text of ISO/IEC 14496-2 MPEG-4 Video FGS VM v 2.0	2926
Text of ISO/IEC 14496-2 Studio Profile Amendment WD 2.0	2927

Description of Core Experiments for MPEG-7 Shape/Motion descriptors	2928
Description of Core Experiments for MPEG-7 Colour/Texture descriptors	2929
Working draft of proposed amendment to ISO/IEC 13818-2	2930
MPEG-7 visual part of XM Version 3	2931
MPEG-4 Video VM 14.0	2932
Report Of The New Formal Verification Tests On MPEG-4 Coding Efficiency for Low and Medium Bit rates	2990
Text of ISO/IEC 14496-4 Version 2 WD 4.0	

2.4.9 to establish the following AdHoc Groups :

Title	Chair(s)	No.	Mtg
AHG on the study of MPEG-2 Video production processes for supplemental information	Gray, McVeigh	2933	Y
AHG on software integration and verification in MPEG-4 video	Ito, Morimatsu	2934	Y
AHG on conformance in MPEG-4 video	Tan	2935	Y
AHG on Object-based content creation for MPEG-7	Kim	2936	Y
AHG on Fine Granularity Scalability in MPEG-4 video	Ohm	2937	Y
AHG on the Studio Profile in MPEG-4 video	Yagasaki	2938	Y
AHG on core experiments for Color/Texture descriptors in MPEG-7	Manjunath, Vinod	2939	Y
AHG on core experiments for Shape/Motion descriptors in MPEG-7	Bober, Jeannin	2940	Y
AHG on editing the documents of the MPEG-4 Visual FDAM and the MPEG-4 video verification model	Jang, Nakaya, Son, Nagumo, Shin, Fukunaga	2941	Y
AHG on editing the document of the MPEG-7 Visual part of XM	Jeannin	2942	Y
AHG on organizing the software integration of MPEG-7 Visual part of XM tools	Herrmann	2943	Y
AHG on MPEG-4 video encoder optimization	Chiang, Sun	2944	Y
AHG on MPEG-7 Generic Description Scheme Development	Salembier, Quackenbush, Saraceno, Jeannin, Walker	2976	Y

2.4.10 to make publicly available the following documents :

Title	No.
Working draft of proposed amendment to ISO/IEC 13818-2	2930

2.5 The Audio group recommends

2.5.1 that the following documents be approved:

Title	No.
DoC on ISO/IEC 13818-4 / FPDAM 3	2912
Text of ISO/IEC 13818-4 / FDAM 3	2913

Study on MPEG-4 Version 1 Audio Conformance FCD	2945
Study on MPEG-4 Version 2 Audio FPDAM	2946
MPEG-4 Version 2 Audio Conformance WD	2947
MPEG-4 Audio conformance work plan	2948
MPEG-7 Audio Core Experiment Methodology	2949
Workplan for MPEG-7 Audio Core Experiment – Sound Effects	2950
Workplan for MPEG-7 Audio Core Experiment – Musical Instruments	2951
Workplan for MPEG-7 Audio Core Experiment – Speech recognition	2952
Workplan for MPEG-4 Version 2 Audio Verification Test	2953
Status and Workplan of MPEG-4 Version 2 Audio Reference Software	2954
Audio FAQ Update	2955
Status and Workplan for MPEG-4 Version 2 Audio Technical Matters	2956
Revised Report on complexity of MPEG-2 AAC Tools	2957

2.5.2 that the following Ad-hoc groups be established:

Title	Chair(s)	No.	Mtg
AHG on MPEG-4 Audio Version 2 Reference Software Editing	H. Purnhagen B. Teichmann	2958	N
AHG on Audio part of MPEG-4 Version 1 and Version 2 Conformance	T. Mlasko T. Moriya	2959	Y
AHG on MPEG-4 Audio Version 2 Study on FPDAM Editing	S-W Kim	2960	Y
AHG on MPEG-7 Audio Core Experiments	Chair: A. Lindsay Co-chair: P. Garner, M Casey, G Peeters, P. Philippe	2961	Y
AHG on MPEG-4 Audio Version 2 Verification Test	R Sperschneider F.Feige	2962	Y
AHG on MPEG-4 Audio Version 2 Technical Matters	B. Grill, M. Iwadare	2963	Y

2.5.3 that the following documents be made publicly available:

Title	No.
Audio FAQ Update	N2955
Audio contribution to Melbourne Press statement	
Revised Report on complexity of MPEG-2 AAC Tools	N2957

2.5.4 The Audio group thanks AT&T for hosting the web site for MPEG-4 conformance bitstreams. (http://www.research.att.com/projects/mpegaudio)

2.5.5 The Audio group thanks Creative Technology Ltd. for supplying Structured Audio Sample Bank Format conformance bitstream.

2.5.6 The audio group thanks the MPEG-4 Platform Verification Bitstream Development Project for supplying MPEG-4 TwinVQ and CELP conformance bitstreams.

2.5.7 The Audio group thanks G. Zoia of EPFL for continuing work on Structured Audio Profiles and Level and for supplying Structured Audio conformance bitstreams.

2.5.8 Audio group supports AHG on MPEG-7 Generic Description Schemes Development (N2976).

2.5.9 The Audio group thanks AT& T, Bosch, Nokia and T-Nova for their efforts in test item selection for the MPEG-4 Version 2 verification test.

2.6 The SNHC group recommends:

2.6.1 that the following documents be approved:

Title	No.
Study of ISO/IEC 14496-2/AMD1 FPDAM	2921

2.6.2 that the following Ad-hoc groups be established:

Title	Chair(s)	No.	Meeting
AHG on study of generic 3D animation	E. Jang, M. Bourges-Sevenier	2982	No
AHG on SNHC conformance	G. Taubin, M. Han, F. Bossen	2983	No
AHG on FBA	E. Petajan, T. Capin	2989	No

2.6.3 that national bodies review and make comments on N2993, Proposed MPEG-4 Version 2 Profiles .

2.6.4 that national bodies voice their support for N2921, Study of 14496-2/AMD1 FPDAM.

2.6.5 that a call for proposals be issued in Maui, if it is shown that tools in MPEG-4 Version 2 are not competitive with tools available outside of MPEG to achieve efficient encoding of generic 3D model animation.

2.6.6 that companies interested in generic 3D model animation make themselves known and actively participate to the activities of the AHG on study of generic 3D animation.

2.6.7 that people involved in FBA activities attend future meetings to ensure the successful completion of MPEG-4 Version 2 activities.

2.6.8 that Mr. Pete Doenges be sincerely thanked for his dedication as chairman of the SNHC group. The group regrets his resignation, and wishes him all the best in his future activities outside of MPEG.

2.6.9 that Dr. Frank Bossen be thanked for successfully chairing the SNHC group during this week.

2.7 The Implementation Study Group recommends:

2.7.1 To approve the following document:

MPEG-7 XM Software Integration: Current status.	N2964

2.7.2 To establish the following ad-hoc group:

Title	Chair(s)	No.	Mtg
AHG on MPEG-7 XM Development	Stephan Herrmann, Mark J. Buxton	N2965	N

2.7.3 To verify the backchannel mechanism for the transmission of SNHC rendering terminal QoS.

2.7.4 That the Audio, Video and MDS groups provide software for the integration of

MPEG-7 XM.

2.7.5 That the following facts are recognized:

1. Using a weighting cost function in a single VBV, VCV and VMV model could be more appropriate than using the current approach with two VCV models.
2. The weighting cost function in a single VBV, VCV and VMV model has insufficiently been validated, compared to the validation of the current approach with two VCV models.

2.8 The Liaison group recommends

2.8.1 the approval of the following responses to National Bodies

Response to National Body Comments	2988

2.8.2 the approval of following people as Liaison representatives.

Organization	Liaison Representatives
TC 46/SC 9	Keith Hill (with Albert Simmonds, present representative)
AGICOA	Didier Mary
OCLC	Jane Hunter
IEC TC100	Kate Grant

2.8.3 the approval of the following Liaison output documents:

Liaison to ITU-R WP 11A	2984
Liaison to SMPTE	2986
Liaison to IETF	2987
Liaison to SMPTE	3008
Liaison to EBU	3009
Liaison to DVB	3010
Liaison to DVD consortium	3011
Liaison to ATSC	3012
Liaison to Pro-MPEG	3013
Liaison to IECTC100	3014
Liaison to SC32/WG2	3015
Liaison to DVB on DSM-CC	3016

2.9 The HoD group

2.9.1 The HODs would like to thank IBM and in particular Peter Schirling for their continued support for MPEG in hosting the World Wide Web site and the ftp sites. The HODs note the increasing administrative burden that results from running the ftp site and web sites. They also note that there is a need for increased security on the part of IBM, and that this will result in changes to the administration of passwords. The HODs therefore invite IBM to indicate the charges that will need to be levied. National Bodies are invited to make proposals for collection of these charges that will result from IBM continuing to run the web and ftp sites for MPEG. They also invite other interested parties to make proposals for alternative arrangements that can deliver the accustomed levels of service. These proposals should be made by the 50th meeting in Maui.

2.9.2 The HoDs recommend the adoption of the following progression of the MPEG-7 work item

2001	Sep		IS
	Jul	23-27	FDIS
	Mar	5-9	FCD
	Jan	15-19	Study of CD
2000	Oct	23-27	CD

2.9.3 **Recommends approval of the meeting plan including**

- **the 50th meeting in Maui, Hawaii, USA over the period 6th to 10th December 1999 and a meeting fee of about US$250.**
- **the 51st meeting in the Netherlands over the period 20th to 24th March 2000 and acceptance of a meeting fee of approximately US$300.**
- **The 52nd meeting in Beijing, China in over the period 17-21 July 2000**
- **the 53rd meeting in France over the period 23-27th October 2000**
- **the 54th meeting in Eilat, Israel over the period 15-19 January 2001**
- **the 55th meeting in Singapore over the period 5-9 March 2001 (to be confirmed).**

2.9.4 **The HODs thank the National Bodies of Singapore and Israel for their kind offers to host future MPEG meetings.**

2.9.5 **The HODs note the US National Body position on allowing software contributions to MPEG for non- normative tools in binary format. The HODs resolve that at present there should be no change to the current policy on this issue.**

2.9.6 **The HoD Group notes the concerns of the French National Body about the protection of intellectual property contributed to the MPEG-7 eXperimental Model (XM). WG11 reaffirms its existing policy on the contribution of software remain in force but that this policy may be subject to further review in the future.**

3 WG11 approves the progression of the following MPEG-2 amendments and corrigenda

DoC on 13818-1 FPDAM 7	**WG11 N2909**
Text o 13818-2 FDAM 6	**WG11 N2911**
Text of 13818-4 FDAM 2	**WG11 N2915**
DoCof 13818-4 FPDAM 2	**WG11 N2914**
DoC on 13818-4/FPDAM 3	**WG11 N2912**
Text of 13818-4/FDAM 3	**WG11 N2913**
DoC on 13818-6 FDAM 1	**WG11 N2916**
Text of 13818-6 FPDAM 1.2	**WG11 N2917**

4 WG11 approves the MPEG-2 workplan:

Part	Title	CfP	WD	CD PDAM PDTR	FCD FPDAM FPDTR	FDIS FDAM DTR DCOR	IS AMD TR COR
MPEG-2							
1/Amd 5	Systems-related table entries for AAC						00/02

1/Amd 6	4:2:2 Profile @High level splice parameters and buffer model for ISO/IEC 13818-7 (AAC)						00/02
1/Amd 7	Transport of ISO/IEC 14496 data over ISO/IEC 13818-1					00/02	00/04
2/Amd 6	Number of lines per frame					00/02	00/04
4/Amd 2	System Target Decoder Model						00/02
4/Amd 3	Audio conformance bitstream					00/03	00/05
6/Amd 1	Additions to support data broadcasting					00/07	00/09
6/Amd 2	Addition to support Synchronized Download Services, Opportunistic Data Services and Resource Announcement in Broadcast and Interactive Services					99/12	00/02

5 WG11 approves the progression of the following MPEG-4 corrigenda

Text of ISO/IEC 14496-1 DCOR 1	**WG11 N3019**
Text of ISO/IEC 14496-2 /DCOR1	**WG11 N2919**
Text of ISO/IEC 14496-6 DCOR 1	**WG11 N3023**

6 WG11 approves the following MPEG-4 amendment

Text of ISO/IEC 14496-5 PDAM 1	**N2918**

7 WG11 approves the following Verification Models

MPEG-4 Systems Ver. 2 VM 8.0	**WG11 N3025**
MPEG-4 Video VM 14.0	**WG11 N2932**

8 WG11 approves MPEG-4 version 4 Conformance WD (WG11 N2991)

9 WG11 approves the MPEG-4 workplan:

Part	Title	CfP	WD	CD PDAM PDTR	FCD FPDAM FPDTR	FDIS FDAM DTR DCOR	IS AMD TR COR
MPEG-4							
1	Systems						99/10
2	Visual						99/11
3	Audio						99/11
4	Conformance Testing					99/12	00/02
5	Reference Software						00/02
6	DMIF						
1/Amd 1	Systems Extensions					99/12	00/02
2/Amd 1	Visual Extensions					99/12	00/02
3/Amd 1	Audio Extensions					99/12	00/02
4/Amd1	Conformance Testing Extensions			99/12	00/07	00/12	01/02

5/Amd 1	Reference Software Extensions			99/12	00/03	00/07	00/09
6/Amd 1	DMIF Extensions					99/12	00/02

10 WG11 approves the MPEG-7 Core Experiments documents (WG11 N2928. 2929, 2972)

11 WG11 approves the DDL document (WG11 N2997).

12 WG11 approves the visual portion of the XM document (WG11 N2931)

13 MPEG-7 approves the draft MPEG-7 development process (WG11 N2999).

14 Software for non-normative parts of MPEG-7 will be offered by contributors of Core Experiments in the form either of source code or of commercial software made available either directly or via a third party to all interested MPEG members for the platform(s) that MPEG will select and for which commercial software must be made available, at conditions acceptable by MPEG, that will be fair and reasonable *(as a point of information, currently the platforms for the XM are Intel/Linux and Intel/Windows32 - but some people use Linux on non-Intel platforms).* Every MPEG member is requested to bring elements and come prepared for a final decision that will be made at the Maui meeting.

15 WG11 approves the MPEG-7 workplan:

Part	Title	CfP	WD	CD PDAM PDTR	FCD FPDAM FPDTR	FDIS DAM DTR DCOR	IS AMD TR COR
MPEG-7							
1	Multimedia Content Description Interface		99/12	00/10	01/03	01/07	01/09

16 WG11 requests the SC29 Secretariat to initiate the procedure to subdivide the MPEG-7 work item in 6 parts:

1. **MPEG-7 Descriptors**
2. **MPEG-7 Description Schemes**
3. **MPEG-7 Description Definition Language**
4. **MPEG-7 Systems**
5. **MPEG-7 Reference SW**
6. **MPEG-7 Conformance.**

The request is made assuming that an satisfactory separation between Descriptors and Description Schemes will be found. If this will not be achieved, parts 1 and 2 will be merged in part 1 and part 2 will remain void.

17 WG11 approves its integrated workplan:

Part	Title	CfP	WD	CD PDAM PDTR	FCD FPDAM FPDTR	FDIS FDAM DTR DCOR	IS AMD TR COR
MPEG-2							
10	Conformance testing extensions - DSM-CC						99/02
1/Amd 5	Systems-related table entries for AAC					99/03	99/05
1/Amd 6	4:2:2 Profile @High level splice parameters and buffer model for ISO/IEC 13818-7 (AAC)					99/03	99/05
1/Amd 7	Transport of ISO/IEC 14496 data over ISO/IEC 13818-1				99/03	99/10	99/12
2/Amd 5	4:2:2 Profile @High Level						99/02
2/Amd 6	Number of lines per frame				99/03	99/10	99/12
4/Amd 2	System Target Decoder Model					99/03	99/05
4/Amd 3	Audio conformance bitstream				99/03	99/10	99/12
6/Amd 1	Additions to support data broadcasting						99/05
6/Amd 2	Addition to support Synchronized Download Services, Opportunistic Data Services and Resource Announcement in Broadcast and Interactive Services			99/03	99/07	99/12	00/02
6/Cor 2						99/03	99/07
MPEG-4							
1	Systems						99/05
2	Visual						99/05
3	Audio						99/05
4	Conformance Testing				99/07	99/12	00/02
5	Reference Software					99/07	99/09
6	DMIF						99/02
1/Amd 1	Systems Extensions			99/03	99/07	99/12	00/02
2/Amd 1	Visual Extensions			99/03	99/07	99/12	00/02
3/Amd 1	Audio Extensions			99/03	99/07	99/12	00/02
4/Amd1	Conformance Testing Extensions			99/12	00/07	00/12	01/02
5/Amd 1	Reference Software Extensions			99/07	99/12	00/03	00/05
6/Amd 1	DMIF Extensions			99/03	99/07	99/12	00/02
MPEG-7							
1	Multimedia Content Description Interface		99/12	00/10	01/03	01/07	01/09

18 WG11 thanks Peter Doenges for his efforts in setting up and chairing the SNHC group since 1995 and regretfully accepts his resignation. Euee S. Jang is appointed as chair of the SNHC group.

19 WG11 approves its revised terms of reference (WG11 N3000)

20 WG11 approves its meeting schedule:

50	**99**	**12**	**06-10**	**Maui, HI**	**US**
51	**00**	**03**	**20-24**	**Noordwijkerhout**	**NL**
52	**00**	**07**	**17-21**	**Beijing**	**CN**
53	**00**	**10**	**23-27**		**FR**
54	**01**	**01**	**15-19**	**Eilat**	**IL**
55	**01**	**03**	**05-09**	**(Singapore)**	**(SG)**

21 WG11 approves its Melbourne press release for publication on the MPEG home page (WG11 N2907).

22 The Convenor would like to express his thanks to Dominique Curet, Jin Woong Kim and Peter Schirling for their support in the preparation of the Friday Plenary.

23 **WG11 would like to thank Standards Australia, the host of its 49th meeting and the sponsors Motorola Research, Sony Australia Pty Ltd, Cable & Wireless – Optus, and Channel 9, in particular David Bruce-Steer, Project Manager and Myra Martin, Helen Gomes and Deralee Vercoe, meeting staff.**

Meeting closed at 21:30

INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11 **N3545**
Beijing, July 2000

Title: **Introduction to MPEG-7 (version 1.0)**
Status: **approved**
Source: **Requirements**
Editor: **José M. Martínez (UPM-GTI, ES)**

TABLE OF CONTENTS

Introduction to MPEG-7

Accessing audio and video used to be a simple matter - simple because of the simplicity of the access mechanisms and because of the poverty of the sources. The transition between two millennia abounds with new ways to produce, offer, filter, search, and manage digitized multimedia information. Broadband is being offered with increasing audio and video quality and speed of access. The trend is clear. In the next few years, users will be confronted with such a large number of content provided by multiple sources that efficient and accurate access to this almost infinite amount of content will seem to be unimaginable. This challenging situation demands a timely solution to the problem. MPEG-7 is the answer to this need.

MPEG-7 aims at offering a comprehensive set of audiovisual description tools to create descriptions, which will form the basis for applications enabling the needed *quality access to content,* which implies good storage solutions, high-performance content identification, proprietary assignation, and fast, ergonomic, accurate and personalized filtering, searching and retrieval. This is a challenging task given the broad spectrum of requirements and targeted multimedia applications, and the broad number of audiovisual features of importance in such context. The question of identifying and managing content is not just restricted to database retrieval applications such as digital libraries, but extends to areas like broadcast channel selection, multimedia editing, and multimedia directory services.

MPEG-7 is an ISO/IEC standard being developed by MPEG (Moving Picture Experts Group), the committee that also developed the Emmy Award winning standards known as MPEG-1 and MPEG-2, and the MPEG-4 standard. The MPEG-1 and MPEG-2 standards are used in many applications, including DVD and digital television.. MPEG-4 provides the standardized technological elements enabling the integration of the production, distribution and content access paradigms of the fields of interactive multimedia, mobile multimedia, interactive graphics and enhanced digital television. This is a kind of repeated in the section ""MPEG-7 and other MPEG standards".

This document gives an introductory overview of the MPEG-7 standard. More information about MPEG-7 can be found at the MPEG-7 website http://drogo.cselt.it/mpeg/ and the MPEG-7 Industry Focus Group website http://www.mpeg-7.com. These web pages contain links to a wealth of information about MPEG, including much about MPEG-7, many publicly available documents, several lists of 'Frequently Asked Questions' and links to other MPEG-7 web pages.

The Context of MPEG-7

More and more audiovisual information is available, from many sources around the world. The information may be represented in various forms of media, such as still pictures, graphics, 3D models, audio, speech, video. Audiovisual information plays an important role in our society, be it recorded in such media as film or magnetic tape or originating, in real time, from some audio or visual sensors and be it analogue or, increasingly, digital. While audio and visual information used to be consumed directly by the human being, there is an increasing number of cases where the audiovisual information is created, exchanged, retrieved, and re-used by computational systems. This may be the case for such scenarios as image understanding (surveillance, intelligent vision, smart cameras, etc.) and media conversion (speech to text, picture to speech, speech to picture, etc.). Other scenarios are information retrieval (quickly and efficiently searching for various types of multimedia documents of interest to the user) and filtering in a stream of audiovisual content description (to receive only those multimedia data items which satisfy the user's preferences). For example, a code in a television program triggers a suitably programmed VCR to record that program, or an image sensor triggers an alarm when a certain visual event happens. Automatic transcoding may be performed from a string of characters to audible information or a search may be performed in a stream of audio or video data. In all these examples, the audiovisual information has been suitably "encoded" to enable a device or a computer code to take some action.

Audiovisual sources will play an increasingly pervasive role in our lives, and there will be a growing need to have these sources processed further. This makes it necessary to develop forms of audiovisual information representation that go beyond the simple waveform or sample-based, compression-based (such as MPEG-1 and MPEG-2) or even objects-based (such as MPEG-4) representations. Forms of representation that allow some degree of interpretation of the information's meaning are necessary. These forms can be passed onto, or accessed by, a device or a computer code. In the examples given above an image sensor may produce visual data not in the form of PCM samples (pixels values) but in the form of objects with associated physical measures and time information. These could then be stored and processed to verify if certain programmed conditions are met. A video recording device could receive descriptions of the audiovisual information associated to a program that would enable it to record, for example, only news with the exclusion of sport. Products from a company could be described in such a way that a machine could respond to unstructured queries from customers making inquiries.

MPEG-7 will be standard for describing the multimedia content data that will support these operational requirements. The requirements apply, in principle, to both real-time and non real-time as well as push and pull applications. MPEG will not standardize or evaluate applications. MPEG may, however, use applications for understanding the requirements and evaluation of technology. It must be made clear that the requirements in this document are derived from analyzing a wide range of potential applications that could use MPEG-7 descriptions. MPEG-7 is not aimed at any one application in particular; rather, the elements that MPEG-7 standardizes shall support as broad a range of applications as possible.

MPEG-7 Objectives

The MPEG-7 standard aims at providing standardized core technologies allowing description of audiovisual data content in multimedia environments. It will extend the limited capabilities of proprietary solutions in identifying content that exist today, notably by including more data types. Audiovisual data content that has MPEG-7 data associated with it, may include: still pictures, graphics, 3D models, audio, speech, video, and composition information about how these elements are combined in a multimedia presentation (scenarios). Special cases of these general data types may include facial expressions and personal characteristics. MPEG-7 description tools do, however, not depend on the ways the described content is coded or stored. It is possible to create an MPEG-7 description of an analogue movie or of a picture that is printed on paper, in the same way as of digitised content.

MPEG-7 Description tools allow to create descriptions (the result of using the MPEG-7 description tools at the users will) of content that may include:

- Information describing the creation and production processes of the content (director, title, short feature movie)
- Information related to the usage of the content (copyright pointers, usage history, broadcast schedule)
- Information of the storage features of the content (storage format, encoding)
- Structural information on spatial, temporal or spatio-temporal components of the content (scene cuts, segmentation in regions, region motion tracking)
- Information about low level features in the content (colors, textures, sound timbres, melody description)
- Conceptual information of the reality captured by the content (objects and events, interactions among objects)

All these descriptions are of course coded in an efficient way for searching, filtering, etc.

To accommodate this variety of complementary content descriptions, MPEG-7 approaches the description of content from several viewpoints. Currently five viewpoints are defined: Creation & Production, Media, Usage, Structural aspects and Conceptual aspects. The five sets of description elements developed on those viewpoints are presented here as separate entities. However, they are interrelated and can be combined in many ways. Depending on the application, some will present and others can be absent or only partly present.

A description generated using MPEG-7 description tools will be associated with the content itself, to allow fast and efficient searching for, and filtering of material that is of interest to the user. The type of content and the query do not have to be the same; for example, visual material may be queried using visual content, music, speech, etc. It is the responsibility of the search engine and filter agent to match the query data to the MPEG-7 description.

Figure 1 explains a hypothetical MPEG-7 chain in practice. The circular boxes depict tools that are doing things, such as encoding or decoding, whereas the square boxes represent static elements, such as a description. The grayed boxes in the figure encompass the normative elements of the MPEG-7 standard. The standard does not describe the process of (automatic) extraction of descriptions/features, nor does it specify the search engine, filter agent, or any other program that can make use of the descriptions.

Note: There can be other streams from content to user; these are not depicted here. Furthermore, it is understood that there might be cases where a binary efficient representation of the description is not needed, and a textual representation would suffice. Thus, the use for the encoder and decoder is optional.

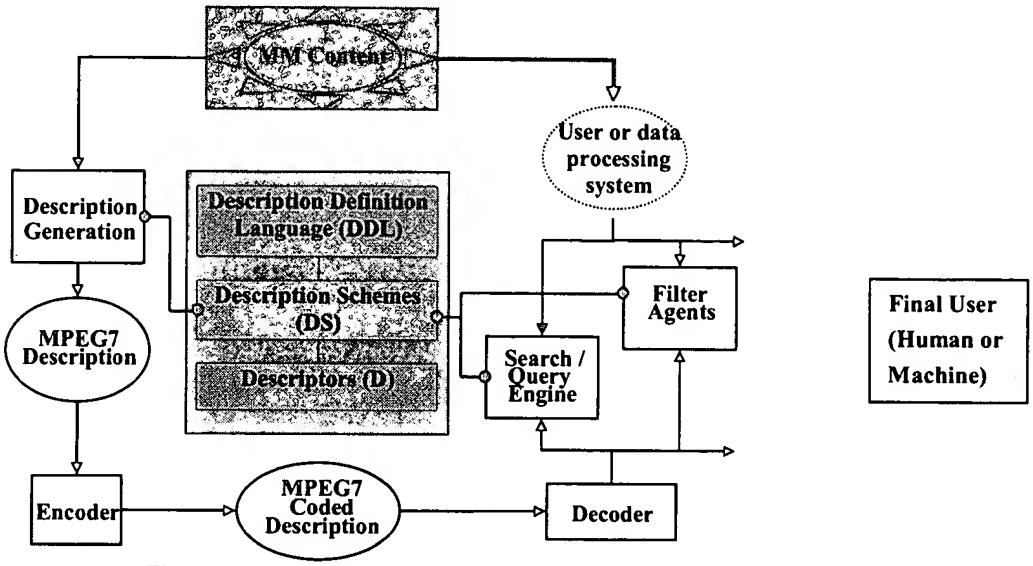


Figure 1: An abstract representation of possible applications using MPEG-7.

MPEG-7 addresses many different applications in many different environments, which means that it needs to provide a flexible and extensible framework for describing audiovisual data. Therefore, MPEG-7 does not define a monolithic system for content description but rather a set of methods and tools for the different viewpoints of the description of audiovisual content. Having this in mind, MPEG-7 is designed to take into account all the viewpoints under consideration by other leading standards such as, SMPTE Metadata Dictionary, Dublin Core, EBU P/Meta, and TV Anytime, which are focused to more specific applications or application domains, whilst MPEG-7 tries to be as generic as possible. MPEG-7 uses also XML Schema as the language of choice for the textual representation of content description and for allowing extensibility of description tools. Considering the popularity of XML, usage of it will facilitate interoperability in the future.

The main elements of the MPEG-7's standard are:

- Descriptors (D): representations of Features, that define the syntax and the semantics of each feature representation,
- Description Schemes (DS), that specify the structure and semantics of the relationships between their components. These components may be both Descriptors and Description Schemes,
- A Description Definition Language (DDL) to allow the creation of new Description Schemes and, possibly, Descriptors and to allows the extension and modification of existing Description Schemes,
- System tools, to support multiplexing of descriptions, synchronization issues, transmission mechanisms, coded representations (both textual and binary formats) for efficient storage and transmission, management and protection of intellectual property in MPEG-7 descriptions, etc.

Creating MPEG-7 Applications

The elements that MPEG-7 standardizes will support a broad range of applications (for example, multimedia digital libraries, broadcast media selection, multimedia editing, home entertainment devices, etc.). MPEG-7 will also make the web as searchable for multimedia content as it is searchable for text today. This would apply especially to large content archives, which are being made accessible to the public, as well as to multimedia catalogues enabling people to identify content for purchase. The information used for content retrieval may also be used by *agents*, for the selection and filtering of broadcasted "push" material or for personalized advertising. Additionally, MPEG-7 descriptions will allow fast and cost-effective usage of the underlying data, by enabling semi-automatic multimedia presentation and editing.

All domains making use of multimedia will benefit from MPEG-7. Considering that at present day it is hard to find one not using multimedia, please extend the list of the examples below using your imagination:

- Digital libraries, Education (image catalogue, musical dictionary, Bio-medical imaging catalogues,...)
- Multimedia editing (personalised electronic news service, media authoring)
- Cultural services (history museums, art galleries, etc.),
- Multimedia directory services (e.g. yellow pages, Tourist information, Geographical information systems)
- Broadcast media selection (radio channel, TV channel,...)
- Journalism (e.g. searching speeches of a certain politician using his name, his voice or his face),
- E-Commerce (personalised advertising, on-line catalogues, directories of e-shops,...)
- Surveillance (traffic control, surface transportation, non-destructive testing in hostile environments, etc.),
- Investigation services (human characteristics recognition, forensics),
- Home Entertainment (systems for the management of personal multimedia collections , including manipulation of content, e.g. home video editing, searching a game, karaoke,...)
- Social (e.g. dating services),

Imagine the things you'll be able to do having MPEG-7 enabled technology. You'll be able to:

- Play a few notes on a keyboard and retrieve a list of musical pieces similar to the required tune, or images matching the notes in a certain way, e.g. in terms of emotions.
- Draw a few lines on a screen and find a set of images containing similar graphics, logos, ideograms,...
- Define objects, including colour patches or textures and retrieve examples among which you select the interesting objects to compose your design.
- On a given set of multimedia objects, describe movements and relations between objects and so search for animations fulfilling the described temporal and spatial relations.
- Describe actions and get a list of scenarios containing such actions.
- Using an excerpt of Pavarotti's voice, obtaining a list of Pavarotti's records, video clips where Pavarotti is singing and photographic material portraying Pavarotti.

Method of Work, Work Plan, and current status

The method of development is comparable to that of the previous MPEG standards. MPEG work is usually carried out in three stages: definition, competition, and collaboration. In the definition phase, the scope, objectives and requirements for MPEG-7 were defined. In the competitive stage, participants worked on their technology by themselves. The end of this stage was marked by the MPEG-7 Evaluation following an open Call for Proposals (CfP). The Call asked for relevant technology fitting the requirements. In answer to the Call, all interested parties, no matter whether they participate or have participated in MPEG, were invited to submit their technology to MPEG. Some 60 parties submitted, in total, almost 400 proposals, after which MPEG made a fair expert comparison between these submissions.

Selected elements of different proposals will be incorporated into a common model (the eXperimentation Model, or XM) during the collaborative phase of the standard. The goal is building the best possible model, which is in essence a draft of the standard itself. During the collaborative phase, the XM is updated and improved in an iterative fashion, until MPEG-7 reaches the Committee Draft (CD) stage in October 2000, after several versions of the Working Draft. Improvements to the XM are made through Core Experiments (CEs). CEs are defined to test the existing tools against new contributions and proposals, within the framework of the XM, according to well-defined test conditions and criteria. Finally, those parts of the XM (or of the Working Draft) that correspond to the normative elements of MPEG-7 will be standardized.

The current work plan for MPEG-7 is shown below:

Call for Proposals	October 1998
Evaluation	February 1999

First version of Working Draft	December1999
Committee Draft	October 2000
Final Committee Draft	February2001
Draft International Standard	July 2001
International Standard	September 2001

Currently MPEG-7 concentrates on the specification of description tools (Descriptors and Description Schemes), together with the development of the MPEG-7 reference software, known as XM (eXperimentation Model). The XML Schema Language was chosen as the base for the Description Definition Language (DDL), but with some further developments since XML Schema is not enough to fulfill all the DDL requirements.

The MPEG-7 Audio group develops a range of Description Tools, from generic audio descriptors (e.g., waveform and spectrum envelopes, fundamental frequency) to more sophisticated description tools like Spoken Content and Timbre. Generic Audio Description tools will allow the search for similar voices, by searching similar envelopes and fundamental frequencies of a voice sample against a database of voices. The Spoken Content Description Scheme (DS) is designed to represent the output of a great number of state of the art Automatic Speech Recognition systems, containing both words and phonemes representations and transition likelihoods. This alleviates the problem of out-of-vocabulary words, allowing retrieval even when the original word was wrongly decoded. The Timbre descriptors (Ds) describe the perceptual features of instrument sound, that make two sounds having the same pitch and loudness appear different to the human ear. These descriptors allow searching for melodies independently of the instruments.

The MPEG-7 Visual group is developing four groups of description tools: Color, Texture, Shape and Motion. Color and Texture Description Tools will allow the search and filtering of visual content (images, graphics, video) by dominant color or textures in some (arbitrarily shaped) regions or the whole image. Shape Description Tools will facilitate "query by sketch" or by contour similarity in image databases, or, for example, searching trademarks in registration databases. Motion Description Tools will allow searching of videos with similar motion patterns that can be applicable to news (e.g. similar movements in a soccer or football game) or to surveillance applications (e.g., detect intrusion as a movement towards the safe zone).

The Multimedia Description Schemes group is developing the description tools dealing with generic (basic structures, common DSs), audiovisual (structure of video and audio) and archival features (collections, streaming). Its central tools deal with content management and content description. Content Management description tools cover the viewpoints of Media, Creation and Production, and Usage. Media description tools allows searching for preferred storage formats, compression qualities, and aspect ratios among others. Creation and Production descriptions tools cover the typical archival and credits information (e.g., title, creators, classification). Usage description tools deal with description related to the use of the described content (e.g. rights, broadcasting dates and places, availability, audience, financial data). The Content Description ones cover both structural and conceptual viewpoints. Structural description tools provide segmentation, both spatial and temporal, of the content. This allows, among other functionalities, assigning descriptions to different regions and segments (e.g., to provide the means for a segment annotation instead of only a global one) and providing importance rating of temporal segments and regions (e.g., allowing to differentiate among regions of the content for adaptive coding with different quality). Conceptual description tools allow providing semantic based description (e.g., linguistic annotations of the content, and object and event description from a knowledge viewpoint). Besides the Content Description and Content Management description tools, there are others targeted to content organization (e.g., to organize an archive of image's descriptions in a repository), navigation and access (e.g., to display a summary of videos through relevant short sequences or keyframes for quick browsing), and user preferences (e.g. for agent based selection or filtering of favorite programs).

The Uniqueness of MPEG-7 in the 21st century Media landscape

How many times have you seen science fiction movies such as 2001, A Space Odyssey or Star Trek and think, 'Wow, we are so far away from having some of the fancy gadgets depicted in these movies!' In 2001, Hal, the talking computer intelligently navigates and retrieves information or runs complex operations instigated by spoken input. What about the communicator in Star Trek? Surely, today's mobile phones are the first signs of a 'Star Trek' communicator where AV content can be broadcasted, filtered, searched, navigated and retrieved.

MPEG-7 is, at last, the beginning of the road to the realization of dreams of so many imaginative minds of the 20th century. MPEG-7 will indeed play a key role to that often heard refrain, 'this is what I thought computers were supposed to do!' MPEG-7 will enable applications that mould computers around human requirements and not humans around computer requirements. Unlike today's state-of-the-art technology, MPEG-7 allows for objective description of features so as to enable content disclosure based on facts, rather than on (unpredictable) human annotations. Finding information by rich spoken queries, hand-drawn images, and humming will improve the user-friendliness of computer systems and finally address what most people expect computers to be able to do.

For professionals, a new generation of applications providing tools for high-quality information search and retrieval will be possible. For example, TV program producers will be able to search with 'laser-precision' for occurrences of famous entities, stored in thousands of hours of audiovisual records, in order to create material for a program about that entity. Program production time will reduce and the quality of program content will increase.

MPEG-7 is a multimedia content description standard, which closely addresses how humans expect to interact with computer systems because it develops rich descriptions that reflect those expectations.

MPEG-7 is about the future of media in the 21st century. This is not an overstatement. MPEG-7 provides a comprehensive and flexible framework for describing the content of multimedia. To describe content implies knowledge of elements it consists of, as well as, knowledge of interrelations between those elements. The most straightforward application is multimedia management, where such knowledge is prerequisite for efficiency and accuracy. However, there are other serious implications. Imprinted knowledge of content and structure, so far elitarian knowledge possessed by content creators only, is made public here, allowing content manipulation, and ultimately content reuse – new content creation. Copyrights issues are not banal here. Other issues and concerns arise, but they are balanced by incredible economical, educational, and ergonomic benefits that will be brought by MPEG-7 technology. Potential concerns will be resolved, and in some years we will not be able to imagine media without MPEG-7 technologies.

References

There are a number of documents available at the MPEG Home Page at http://drogo.cselt.it/mpeg/, including:

- MPEG-7 Requirements
- MPEG-7 Applications
- MPEG-7 Context, Objectives and Technical Roadmap
- MPEG-7 Principal Concepts List
- MPEG-7 Overview
- MPEG-7 DDL WD
- MPEG-7 Visual WD
- MPEG-7 Audio WD
- MPEG-7 MDS WD and XM

Information more focused to industry is also available at the MPEG-7 Industry Focus Group Web site at http://www.mpeg-7.com.

Contact points

- Neil Day (neil@garage.co.jp)

For Technical Issues

- Requirements: Rob Koenen (r.h.koenen@research.kpn.com)
- Audio: Adam Lindsay (atl@comp.lancs.ac.uk)
- Visual: Thomas Sikora (sikora@hhi.de)
- Multimedia DS: Philippe Salembier (philippe@gps.tsc.upc.es)
- Systems: Olivier Avaro (olivier.avaro@francetelecom.fr)
- XM Software Implementation: Stephan Herrmann (stephanh@lis.e-technik.tu-muenchen.de)

Editor:

- José M. Martínez (jms@gti.ssr.upm.es)

Annex A: MPEG-7 Terminology

1. **Data**

Definition

Data is audiovisual information that will be described using MPEG-7, regardless of storage, coding, display, transmission, medium, or technology.

Notes

This definition is intended to be sufficiently broad to encompass graphics, still images, video, film, music, speech, sounds, text and any other relevant AV medium.

Examples

Examples for MPEG-7 data are an MPEG-4 stream, a video tape, a CD containing music, sound or speech, a picture printed on paper, and an interactive multimedia installation on the web.

2. **Feature**

Definition

A Feature is a distinctive characteristic of the data which signifies something to somebody.

Notes

Features themselves cannot be compared without a meaningful feature representation (descriptor) and its instantiation (descriptor value) for a given data set.

Examples

Some examples are: color of an image, pitch of a speech segment, rhythm of an audio segment, camera motion in a video, style of a video, the title of a movie, the actors in a movie etc.

3. **Descriptor**

Definition

A Descriptor (D) is a representation of a Feature. A Descriptor defines the syntax and the semantics of the Feature representation.

Notes

A descriptor allows an evaluation of the corresponding feature via the descriptor value. It is possible to have several descriptors representing a single feature, i.e. to address different relevant requirements.

Examples

For example for the color feature, possible descriptors are: the color histogram, the average of the frequency components, the motion field, the text of the title, etc. More examples of Features and their associated Descriptors are provided in Table 1.

4. **Descriptor Value**

Definition

A Descriptor Value is an instantiation of a Descriptor for a given data set (or subset thereof).

Notes

Descriptor Values are combined via the mechanism of a Description Scheme (see point 5) to form a Description (see point 6).

Examples

-

5. **Description Scheme**

Definition

A Description Scheme (DS) specifies the structure and semantics of the relationships between its components, which may be both Descriptors and Description Schemes.

Notes

The distinction between a DS and a D is, that a D contains only basic data types, as provided by the DDL (see point 8), and does not refer to another D or (sub)DS.

Examples

A movie, temporally structured as scenes and shots, including some textual descriptors at the scene level, and color, motion and some audio descriptors at the shot level.

6. **Description**

   Definition

   A Description consists of a DS (structure) and the set of Descriptor Values (instantiations) that describe the Data.

   Notes

   Depending on the completeness of the set of Descriptor Values, the DS may be fully or partially instantiated. Whether or not the DS is actually present in the Description depends on technical solutions still to be provided.

   Examples

   -

7. **Coded Description**

   Definition

   A Coded Description is a Description that has been encoded to fulfil relevant requirements such as compression efficiency, error resilience, random access, etc.

   Notes

   -

   Examples

   -

8. **Description Definition Language**

   Definition

   The Description Definition Language (DDL) is a language that allows the creation of new Description Schemes and, possibly, Descriptors. It also allows the extension and modification of existing Description Schemes.

   Notes

   It is not yet clear to which extend the DDL will allow the creation of new descriptors.

   Examples

   -

Annex B: MPEG-7 FAQs

1. **What is MPEG-7?**

   MPEG-7 will be a standardised description of various types of multimedia information. This description will be associated with the content itself, to allow fast and efficient searching for material that is of interest to the user. MPEG-7 is formally called 'Multimedia Content Description Interface'. The standard does not comprise the (automatic) extraction of descriptions/features. Nor does it specify the search engine (or any other program) that can make use of the description.

2. **From whom or where did the demand for MPEG-7 come?**

   The demand logically follows the increasing availability of digital audiovisual content. MPEG members recognised this demand, and initiated a new work item. The work on the definition of MPEG-7 has already started to attract new people to MPEG.

3. **Why is MPEG-7 needed?**

   Nowadays, more and more audiovisual information is available, from many sources around the world. Also, there are people who want to use this audiovisual information for various purposes. However, before the information can be used, it must be located. At the same time, the increasing availability of potentially interesting material makes this search more difficult. This challenging situation led to the need of a solution to the problem of quickly and efficiently searching for various types of multimedia material interesting to the user. MPEG-7 wants to answer to this need, providing this solution.

4. **Who is currently participating in the development of the MPEG-7 standard?**

The people taking part in defining MPEG-7 represent broadcasters, equipment manufacturers, digital content creators and managers, transmission providers, publishers and intellectual property rights managers, as well as university researchers.

5. **Where are you in the process of specifying the MPEG-7 standard?**

We are in the collaborative phase of the standardisation process. This means that we have passed the Call for Proposals and the evaluation of the submissions to that CfP. We are currently performing experiments (so-called Core Experiments) to continuously improve the technology on the table for standardization. This testing is carried out in a common environment, called the eXperimentation Model (XM). Experiments are carried out in well-defined test conditions and according to pre-defined criteria. The goal is to develop the best possible standard.

6. **Will MPEG-7 include audio or video content recognition?**

The standardisation of audiovisual content recognition tools is beyond the scope of MPEG-7. Following its principle 'specifying the minimum for maximum usability, MPEG-7 will concentrate on standardising a representation that can be used for description. Development of audiovisual content recognition tools will be a task for industries which will build and sell MPEG-7 enabled products. In developing the standard, however, MPEG might build some coding tools, just as it did with the predecessors of MPEG-7, namely MPEG-1, -2 and -4. Also for these standards, coding tools were built for research purposes, but they did not become part of the standard itself.

7. **Will MPEG-7 support audio or video content retrieval?**

In the same way that MPEG will not standardise the tools to generate the description, MPEG-7 will also not standardise the tools that use the description. It might however be necessary to address the interface between the description and the search engine.

8. **What form will the "descriptions" of multimedia content in MPEG-7 take?**

The words 'descriptions' or 'features' represent a rich concept, that can be related to several levels of abstraction. Descriptions vary according to the types of data. Furthermore, different types of descriptions are necessary for different purposes of the categorisation.

9. **Will the standard allow automatic extraction of descriptions as well as manual entry?**

The descriptions that conform to the MPEG-7 standard could be entered by hand, but they could also be automatically extracted. Some features can be best extracted automatically (colour, texture), but for some other features ('this scene contains three shoes and that music was recorded in 1995') this is very hard or even impossible.

10. **A 'Call for Proposals', how does that work?**

A Call for Proposals (CfP) asks for technology for inclusion in the standard. It is addressed at all interested parties, no matter whether they participate or have participated in MPEG. MPEG work is usually carried out in two stages, a competitive and a collaborative one. In the competitive stage, participants work on their technology by themselves. In answer to the CfP, people submit their technology to MPEG, after which MPEG makes a fair comparison between the submissions. Based on the outcome of the evaluation, MPEG decided which proposals to use for the collaborative stage. In this stage, members of the Experts Group work together on improving and expanding the standard under construction, building on the selected proposals.

11. **What is the relationship between MPEG-7 and other MPEG activities?**

MPEG-7 can be used independently of the other MPEG standards - the description might even be attached to an analog movie. The representation that is defined within MPEG-4, i.e. the representation of audiovisual data in terms of objects, is however very well suited to what will be built on the MPEG-7 standard. This representation is basic to the process of categorisation. In addition, MPEG-7 descriptions could be used to improve the functionality of previous MPEG standards.

12. **If I want to get involved in MPEG-7, what do I need to know about the other MPEG standards?**

In principle, knowledge about the other three MPEG standards is not required for taking part in the MPEG-7 work. However, since some of MPEG-7's tools may be close to those of MPEG-4, some knowledge about them could be useful.

13. If I want to know more about the other MPEG standards, where do I look?

You can start by taking a look at MPEG's home page (http://www.cselt.it/mpeg/) which contains many useful references, including more lists with "Frequently Asked Questions" about MPEG activities.

14. So what happened to MPEG-5 and -6? (And how about 3?)

MPEG-3 existed once upon a time, but its goal, enabling HDTV, could be accomplished using the tools of MPEG-2, and hence the work item was abandoned. So after 1,2 and 4, there was much speculation about the next number. Should it be 5 (the next) or 8 (creating an obvious binary pattern)? MPEG, however, decided not to follow either logical expansion of the sequence, but chose the number of 7 instead. So MPEG-5 and MPEG-6 are, just like MPEG-3, not defined.

15. When will MPEG-7 replace the existing MPEG-1 and MPEG-2 standards?

MPEG-7 will not replace MPEG-1, MPEG-2 or MPEG-4. It is intended to provide complementary functionality to these other MPEG standards: representing information about the content, not the content itself ("the bits about the bits". This functionality is the standardisation of multimedia content descriptions.

17. If I want to know more about, be involved in, or give an input to the MPEG-7 development process, whom should I contact?

You can contact any of the people listed below with their email addresses and telephone numbers. To visit MPEG meetings you need to be on your national delegation, but the people listed in the contact points can explain how this works.

Annex C: MPEG-7 and other MPEG Standards

Currently there are three MPEG standards dealing with compression, decompression, processing, and coded representation of moving pictures, audio and their combination.

MPEG-1 is a standard for storage and retrieval of moving pictures and audio on storage media, that is very successful. It is the de-facto form of storing moving pictures and audio on the World Wide Web and is used in millions of Video CDs. Digital Audio Broadcasting (DAB) is a new consumer market that makes use of MPEG-1 audio coding.

MPEG-2 is a standard for digital television, that has been the timely response for the satellite broadcasting and cable television industries in their transition from analogue to digital formats. Millions of set-top boxes incorporating MPEG-2 decoders have been sold in the last 3 years.

MPEG-4 is a standard for multimedia applications that supports the creation of rich, reusable, and interactive multimedia content that can be used by different distribution networks (broadcasting, internet, CDs,...) and terminals (PCs with Web browsers, TV sets, Set-Top-Boxes, DVD players, ...). MPEG-4 is the first real multimedia representation standard, allowing interactivity and a combination of natural and synthetic materials, coded in the form of objects that are integrated to compose multimedia presentations (scenarios).

In principle, MPEG-1, -2, and -4 are designed to represent the information itself, while MPEG-7 is meant to represent information about the information. Looking from another perspective: MPEG-1, -2, and -4 make content available, while MPEG-7 allows you to find the content you need.

MPEG-7 can be used independently of the other MPEG standards - the description might even be attached to an analog movie. MPEG-7 descriptions could be used to improve the functionalities of previous MPEG standards, but will not replace MPEG-1, MPEG-2 or MPEG-4. It is intended to provide complementary functionality to these other MPEG standards: representing information about the content, not the content itself ("the bits about the bits").

Besides these standards, MPEG started recently the development of MPEG-21, a standard that aims at creating a Multimedia Framework taking into consideration the different components involved in the delivery of content from the creator to the user.

INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11 N2460
MPEG98
October 1998 / Atlantic City, USA

Source: Requirements Group
Status: Approved
Title: MPEG-7: Context and Objectives (version - 10 Atlantic City)

MPEG-7 Context and Objectives

1. Context

More and more audio-visual information is available in digital form, in various places around the world. Along with the information, people appear that want to use it. Before one can use any information, however, it will have to be located first. At the same time, the increasing availability of potentially interesting material makes this search harder. Currently, solutions exist that allow searching for textual information. Many text-based search engines are available on the World Wide Web, and they are among the most visited sites - indicating they foresee a real demand. Identifying information is, however, not possible for audio-visual content, as no generally recognised description of this material exists. In general, it is not possible to efficiently search the web for, say, a picture of 'the Motorbike from Terminator II', or to search a sequence where "King Lear congratulates his assistants on the night after the battle," or to search for "twenty minutes of video according to my preferences of today". In specific cases, solutions do exist. Multimedia databases on the market today allow

searching for pictures using characteristics like colour, texture and information about the shape of objects in the picture. One could envisage a similar example for audio, in which one can whistle a melody to find a song.
The question of finding content is not restricted to database retrieval applications; also in other areas similar questions exist. For instance, there is an increasing amount of (digital) broadcast channels available, and this makes it harder to select the broadcast channel (radio or TV) that is potentially interesting.

2. MPEG-7 Objectives

In October 1996, MPEG started a new work item to provide a solution to the questions described above. The new member of the MPEG family, called "Multimedia Content Description Interface" (in short 'MPEG-7'), will extend the limited capabilities of proprietary solutions in identifying content that exist today, notably by including more data types. In other words: MPEG-7 will specify a standard set of descriptors that can be used to describe various types of multimedia information. MPEG-7 will also standardise ways to define other descriptors as well as structures (Description Schemes) for the descriptors and their relationships (see also 2.1 What is the Scope of the Standard). This description (i.e. the combination of descriptors and description schemes) shall be associated with the content itself, to allow fast and efficient searching for material of a user's interest. MPEG-7 will also standardise a language to specify description schemes, i.e. a Description Definition Language (DDL). AV material that has MPEG-7 data associated with it, can be indexed and searched for. This 'material' may include: still pictures, graphics, 3D models, audio, speech, video, and information about how these elements are combined in a multimedia presentation ('scenarios', composition information). Special cases of these general data types may include facial expressions and personal characteristics.

MPEG-7, like the other members of the MPEG family, is a standard representation of audio-visual information satisfying particular requirements. The MPEG-7 standard builds on other (standard) representations such as analogue, PCM, MPEG-1, -2 and -4. One functionality of the standard is to provide references to suitable portions of them. For example, perhaps a shape descriptor used in MPEG-4 is useful in an MPEG-7 context as well, and the same may apply to motion vector fields used in MPEG-1 and MPEG-2.
MPEG-7 descriptors do, however, not depend on the ways the described content is coded or stored. It is possible to attach an MPEG-7 description to an analogue movie or to a picture that is printed on paper. Even though the MPEG-7 description does not depend on the (coded) representation of the material, the standard in a way builds on MPEG-4, which provides the means to encode audio-visual material as objects having certain relations in time (synchronisation) and space (on the screen for video, or in the room for audio). Using MPEG-4 encoding, it will be possible to attach descriptions to elements (objects) *within* the scene, such as audio and visual objects.. MPEG-7 will allow different granularity in its descriptions, offering the possibility to have different levels of discrimination.

Because the descriptive features must be meaningful in the context of the application,

they will be different for different user domains and different applications.
This implies that the same material can be described using different types of features, tuned to the area of application. To take the example of visual material: a lower abstraction level would be a description of e.g. shape, size, texture, colour, movement (trajectory) and position ('where in the scene can the object be found?). And for audio: key, mood, tempo, tempo changes, position in sound space. The highest level would give semantic information: 'This is a scene with a barking brown dog on the left and a blue ball that falls down on the right, with the sound of passing cars in the background.' All these descriptions are of course coded in an efficient way - efficient for search that is. Intermediate levels of abstraction may also exist.

The level of abstraction is related to the way the features can be extracted: many low-level features can be extracted in fully automatic ways, whereas high level features need (much) more human interaction.

Next to having a description of the content, it may also be required to include other types of information about the multimedia data:

- *The form* - An example of the form is the coding scheme used (e.g. JPEG, MPEG-2), or the overall data size. This information helps determining whether the material can be 'read' by the user.
- *Conditions for accessing the material* - This could include copyright information, and price;
- *Classification* - This could include parental rating, and content classification into a number of pre-defined categories;
- *Links to other relevant material* - The information may help the user speeding up the search.
- *The context* - In the case of recorded non-fiction content, it is very important to know the occasion of the recording (e.g. Olympic Games 1996, final of 200 meter hurdles, men)

In many cases, it will be desirable to use textual information for the descriptions. Care must be taken, however, that the usefulness of the descriptions is as independent from the language area as possible. (A very clear example where text comes in handy is in giving names of authors, film, places.)

MPEG-7 data may be physically located with the associated AV material, in the same data stream or on the same storage system, but the descriptions could also live somewhere else on the globe. When the content and its descriptions are not co-located, mechanisms that link AV material and their MPEG-7 descriptions are useful; these links should work in both directions.

2.1 What is the Scope of the Standard?

MPEG-7 will address applications that can be stored (on-line or off-line) or streamed (e.g. broadcast, push models on the Internet), and can operate in both real-time and non real-time environments. A 'real-time environment' means that information is associated with the content while it is being captured.

Figure 1 below shows a highly abstract block diagram of a possible MPEG-7 processing chain, included here to explain the scope of the MPEG-7 standard. This chain includes feature extraction (analysis), the description itself, and the search engine (application). To fully exploit the possibilities of MPEG-7 descriptions, automatic extraction of features (or 'descriptors') will be extremely useful. It is also clear that automatic extraction is not always possible, however. As was noted above, the higher the level of abstraction, the more difficult automatic extraction is, and interactive extraction tools will be of good use. However useful they are, neither automatic nor semi-automatic feature extraction algorithms will be inside the scope of the standard. The main reason is that their standardisation is not required to allow interoperability, while leaving space for industry competition. Another reason not to standardise analysis is to allow making good use of the expected improvements in these technical areas.

Also the search engines will not be specified within the scope of MPEG-7; again this is not necessary, and here too, competition will produce the best results.

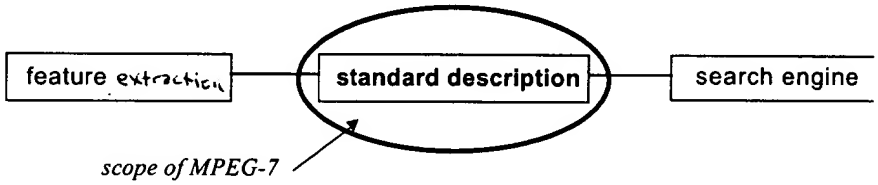


Figure 1: Scope of MPEG-7

To provide a better understanding of the introduced terminology, i.e. Descriptor, Description Scheme, and DDL, please find below Figures 2 – 4. The dotted boxes in the figures encompasses the normative elements of the MPEG-7 standard. Note that the presence of a box or ellipse in one of this drawings does not imply that the corresponding element shall be present in all MPEG-7 applications.

Figure 2 shows the extensibility of the above concepts. Note, the arrows from DDL to DS signify that the DSs are generated using DDL. Furthermore, the drawing reveals the fact that you can build a new DS using an existing DS.

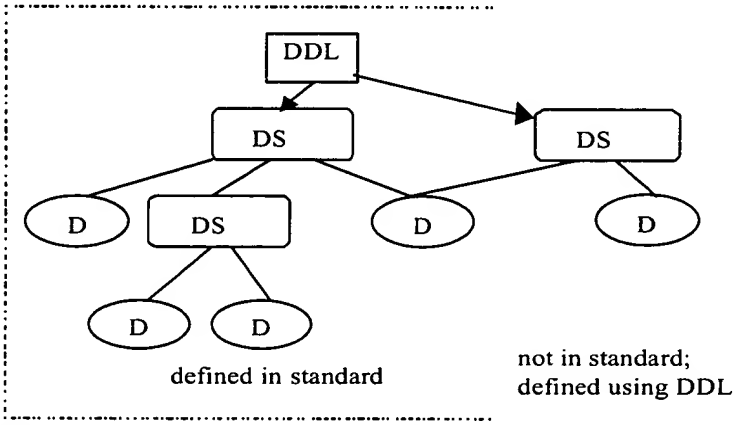


Figure 2: An abstract representation of possible relations between Ds and DSs.

Figure 3 that the DDL provides the mechanism to built a description scheme which in turn forms the basis for the generation of a description. The instantiation of the DS is

described as part of Figure 4.

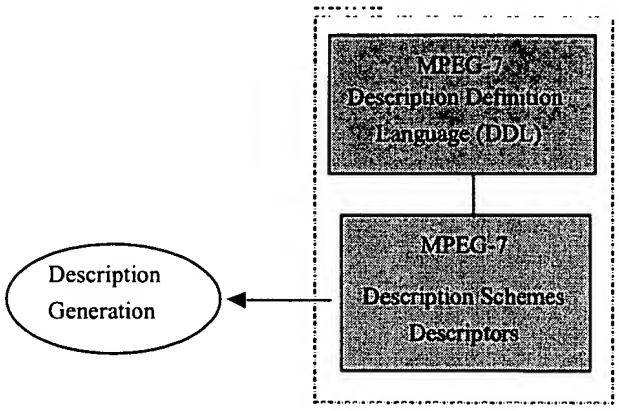


Figure 3: The role of Ds and DSs for the generation of descriptions

Figure 4 explains how MPEG-7 would work in practice. Note: There can be other streams from content to user; these are not depicted here. Furthermore, the use for the encoder and decoder is optional.

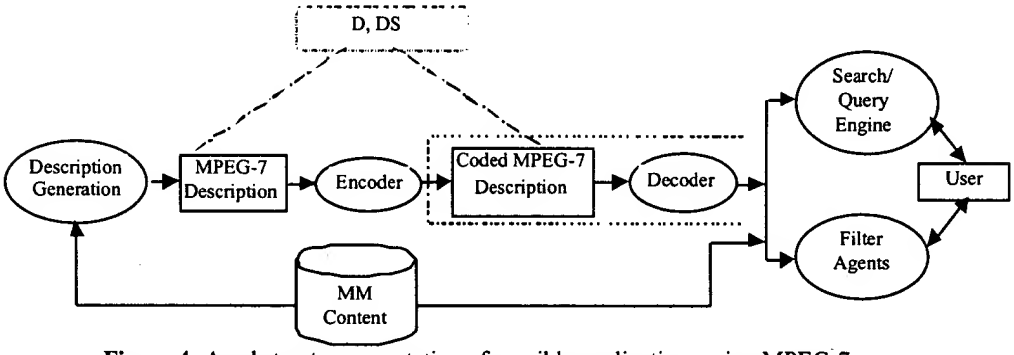


Figure 4: An abstract representation of possible applications using MPEG-7.

The emphasis of MPEG-7 will be the provision of novel solutions for audio-visual content description. Thus, addressing text-only documents will not be among the goals of MPEG-7. However, audio-visual content may include or refer to text in addition to its audio-visual information. MPEG-7, , therefore, will consider existing solutions developed by other standardisation organisations for text only documents and support them as appropriate.

Besides the descriptors themselves, the database structure plays a crucial role in the final retrieval's performance. To allow the desired fast judgement about whether the material is of interest, the indexing information will have to be structured, e.g. in a hierarchical or associative way.

More detailed descriptions of requirements can be found in the 'MPEG-7 Requirements Document' [1].

3. Areas of Interest

There are many applications and application domains which will benefit from the MPEG-7 standard. A few application examples are:

- Digital libraries (image catalogue, musical dictionary,...)
- Multimedia directory services (e.g. yellow pages)
- Broadcast media selection (radio channel, TV channel,...)
- Multimedia editing (personalised electronic news service, media authoring)

The potential applications are spread over the following application domains:

- Education,
- Journalism (e.g. searching speeches of a certain politician using his name, his voice or his face),
- Tourist information,
- Cultural services (history museums, art galleries, etc.),
- Entertainment (e.g. searching a game, karaoke),
- Investigation services (human characteristics recognition, forensics),
- Geographical information systems,
- Remote sensing (cartography, ecology, natural resources management, etc.),
- Surveillance (traffic control, surface transportation, non-destructive testing in hostile environments, etc.),
- Bio-medical applications,
- Shopping (e.g. searching for clothes that you like),
- Architecture, real estate, and interior design,
- Social (e.g. dating services), and
- Film, Video and Radio archives.

The way MPEG-7 data will be used to answer user queries is outside the scope of the standard. In principle, any type of AV material may be retrieved by means of any type of query material. This means, for example, that video material may be queried using video, music, speech, etc. It is to the search engine to match the query data and the MPEG-7 AV description. A few query examples are:

1. *Music*

Play a few notes on a keyboard and get in return a list of musical pieces containing (or close to) the required tune or images somehow matching the notes, e.g. in terms of emotions.

2. *Graphics*

Draw a few lines on a screen and get in return a set of images containing similar graphics, logos, ideograms,...

3. *Image*

Define objects, including colour patches or textures and get in return examples among which you select the interesting objects to compose your image.

4. *Movement*

On a given set of objects, describe movements and relations between objects and get in return a list of animations fulfilling the described temporal and spatial relations.

5. *Scenario*

On a given content, describe actions and get a list of scenarios where similar actions happen.

5. *Voice*
Using an excerpt of Pavarotti's voice, and getting a list of Pavarotti's records, video clips where Pavarotti is singing or video clips where Pavarotti is present.

More detailed descriptions of applications can be found in the 'MPEG-7 Applications Document' [2].

4. Method of Work and Work Plan

The method of development is comparable to that of the previous MPEG standards. After defining the requirements (this process has already started), an open Call for Proposals will be issued. The Call will ask for relevant technology fitting the requirements, and after an evaluation of the technology that was received, a choice will be made and development will continue with the most promising submission(s). In the course of developing the standard, additional calls can be issued when not enough technology is present within MPEG to meet the requirements, and there is a reasonable belief that the technology does indeed exist.

As this new MPEG work item will require technology available in technological areas not yet sufficiently represented in the MPEG community, it shall be necessary to seek the collaboration of new experts in the relevant areas. As always, MPEG is open to anyone interested to participate and contribute.

The preliminary work plan for MPEG-7 foresees:

Call for Proposals	October 1998
Working Draft	December1999
Committee Draft	October 2000
Final Committee Draft	February2001
Draft International Standard	July 2001
International Standard	September 2001

More detailed regarding the call for proposals can be found in the 'MPEG-7 Evaluation Document' [3] and the 'MPEG-7 Proposal Package Description (PPD)' [4].

5. Frequently Asked Questions

1. What is MPEG-7?

MPEG-7 will be a standardised description of various types of multimedia information. This description will be associated with the content itself, to allow fast and efficient searching for material that is of interest to the user. MPEG-7 is formally called 'Multimedia Content Description Interface'.

The standard does not comprise the (automatic) extraction of descriptions/features. Nor does it specify the search engine (or any other program) that can make use of the description.

2. From whom or where did the demand for MPEG-7 come?

The demand logically follows the increasing availability of digital audio-visual content. MPEG members recognised this demand, and initiated a new work item. The work on the definition of MPEG-7 has already started to attract new people to MPEG.

3. Why is MPEG-7 needed?

Nowadays, more and more audio-visual information is available, from many sources around the world. Also, there are people who want to use this audio-visual information for various purposes. However, before the information can be used, it must be located. At the same time, the increasing availability of potentially interesting material makes this search more difficult. This challenging situation led to the need of a solution to the problem of quickly and efficiently searching for various types of multimedia material interesting to the user. MPEG-7 wants to answer to this need, providing this solution.

4. Who is currently participating in the development of the MPEG-7 standard?

The people taking part in defining MPEG-7 represent broadcasters, equipment manufacturers, digital content creators and managers, transmission providers, publishers and intellectual property rights managers, as well as university researchers.

5. Where are you in the process of specifying the MPEG-7 standard?

We are in the phase of defining the scope of the standard and its requirements, and the ideas are likely to evolve considerably. Much is still open to input from interested parties, and MPEG is aware that useful work has already been carried out in several areas. The work plan is as follows:

Call for Proposals	*October 1998*
Working Draft	*December1999*
Committee Draft	*October 2000*
Final Committee Draft	*February2001*
Draft International Standard	*July 2001*
International Standard	*September 2001*

6. Will MPEG-7 include audio or video content recognition?

The standardisation of audio-visual content recognition tools is beyond the scope of MPEG-7. Following its principle 'specifying the minimum for maximum usability, MPEG-7 will concentrate on standardising a representation that can be used for description. Development of audio-visual content recognition tools will be a task for industries which will build and sell MPEG-7 enabled products.

In developing the standard, however, MPEG might build some coding tools, just as it did with the predecessors of MPEG-7, namely MPEG-1, -2 and -4. Also for

these standards, coding tools were built for research purposes, but they did not become part of the standard itself.

7. Will MPEG-7 support audio or video content retrieval?

In the same way that MPEG will not standardise the tools to generate the description, MPEG-7 will also not standardise the tools that use the description. It might however be necessary to address the interface between the description and the search engine.

8. What form will the "descriptions" of multimedia content in MPEG-7 take?

The words 'descriptions' or 'features' represent a rich concept, that can be related to several levels of abstraction. Descriptions vary according to the types of data. Furthermore, different types of descriptions are necessary for different purposes of the categorisation.

9. Will the standard allow automatic extraction of descriptions as well as manual entry?

The descriptions that conform to the MPEG-7 standard could be entered by hand, but they could also be automatically extracted. Some features can be best extracted automatically (colour, texture), but for some other features ('this scene contains three shoes and that music was recorded in 1995') this is very hard or even impossible.

10. A 'Call for Proposals', how does that work?

A Call for Proposals (CfP) asks for technology for inclusion in the standard. It is addressed at all interested parties, no matter whether they participate or have participated in MPEG.

MPEG work is usually carried out in two stages, a competitive and a collaborative one. In the competitive stage, participants work on their technology by themselves. In answer to the CfP, people submit their technology to MPEG, after which MPEG makes a fair comparison between the submissions. In MPEG-2 and -4 this was done using subjective tests and additional expert evaluation. How such evaluations will be carried out for MPEG-7 is not yet known, but this will be described in the CfP when it is published in 1998.

Based on the outcome of the evaluation, MPEG will decide which proposals to use for the collaborative stage. In this stage, members of the Experts Group work together on improving and expanding the standard under construction, building on the selected proposals.

Before the final CfP in November 1998, preliminary versions may be published. This is comparable to what happened for MPEG-4.

11. What is the relationship between MPEG-7 and other MPEG activities?

MPEG-7 can be used independently of the other MPEG standards - the description might even be attached to an analog movie. The representation that is defined within MPEG-4, i.e. the representation of audio-visual data in terms of objects, is however very well suited to what will be built on the MPEG-7 standard. This

representation is basic to the process of categorisation. In addition, MPEG-7 descriptions could be used to improve the functionality of previous MPEG standards.

12. If I want to get involved in MPEG-7, what do I need to know about the other MPEG standards?

In principle, knowledge about the other three MPEG standards is not required for taking part in the MPEG-7 work. However, since some of MPEG-7's tools may be close to those of MPEG-4, some knowledge about them could be useful.

13. If I want to know more about the other MPEG standards, where do I look?

You can start by taking a look at MPEG's home page (http://www.cselt.it/mpeg/) which contains many useful references, including more lists with "Frequently Asked Questions" about MPEG activities.

14. So what happened to MPEG-5 and -6? (And how about 3?)

MPEG-3 existed once upon a time, but its goal, enabling HDTV, could be accomplished using the tools of MPEG-2, and hence the work item was abandoned. So after 1,2 and 4, there was much speculation about the next number. Should it be 5 (the next) or 8 (creating an obvious binary pattern)? MPEG, however, decided not to follow either logical expansion of the sequence, but chose the number of 7 instead. So MPEG-5 and MPEG-6 are, just like MPEG-3, not defined.

15. When will MPEG-7 replace the existing MPEG-1 and MPEG-2 standards?

MPEG-7 will not replace MPEG-1 MPEG-2 or in fact MPEG-4 it is intended to provide complementary functionality to these other MPEG standards: representing information about the content, not the content itself ("the bits about the bits") This functionality is the standardisation of multimedia content descriptions.

16. If I want to know more about, be involved in, or give an input to the MPEG-7 development process, whom should I contact?

You can contact any of the people listed below with their email addresses and telephone numbers. To visit MPEG meetings you need to be on your national delegation, but the people listed below can explain how this works.

Rob Koenen (KPN Research - the Netherlands / chairman MPEG Requirements)
r.h.koenen@research.kpn.com +31 70 332 5310

Sylvie Jeannin (Philips Research – US / evaluation contact)
sjn@philabs.research.philips.com

Fernando Pereira (Instituto Superior Técnico - Portugal)
fp@lx.it.pt +351 1 8418460

Ibrahim Sezan (Sharp Labs - USA)
sezan@sharplabs.com +1 360 817 8401

Adam Lindsay (Riverland, Belgium/ audio contact)
adam@riv.be + 32 2 721 5454

Frank Nack (GMD-IPSI – Germany / requirements contact)
nack@darmstadt.gmd.de + 49 6151 869833

Seungyup Paek (Columbia University – US / test material contact)
Syp@ee.columbia.edu +1 212 854 7447

V.V.Vinod (Kent Ridge digital Labs – Singapore / PPD contact)
vinod@krdl.org.sg +65 874 5225

References

[1] *MPEG Requirements Group, "MPEG-7Requirements Document", Doc. ISO/MPEG N2461, MPEG Atlantic City Meeting, October 1998*

[2] *MPEG Requirements Group, "Applications for MPEG-7", Doc. ISO/MPEG N2462, MPEG Atlantic City Meeting, October 1998*

[3] *MPEG Requirements Group, " MPEG-7Evaluation Procedure", Doc. ISO/MPEG N2463, MPEG Atlantic City Meeting, October 1998*

[4] *MPEG Requirements Group,* "MPEG-7 Proposal Package Description (PPD)", *Doc. ISO/MPEG N2464, MPEG Atlantic City Meeting, October 1998*

STORIES

D-Lib Magazine
September 1999

Volume 5 Number 9

ISSN 1082-9873

MPEG-7
Behind the Scenes

Jane Hunter
Distributed Systems Technology Centre
University of Queensland
jane@dstc.edu.au

Abstract

The purpose of this article is to provide a better understanding of the objectives and components of the MPEG-7, "Multimedia Content Description Interface" standard, an overview of the current state of its development and an idea of its expected impact on digital libraries of the future.

Introduction

It's clearly much more fun to develop multimedia content than to index it. The amount of multimedia content available -- in digital archives, on the World Wide Web, in broadcast data streams and in personal and professional databases -- is growing out of control. But this enthusiasm has led to increasing difficulties in accessing, identifying and managing such resources due to their volume and complexity and a lack of adequate indexing standards. The large number of recently-funded DLI-2 projects related to the resource discovery of different media types, including music, speech, video and images, indicates an acknowledgement of this problem and the importance of this field of research for digital libraries. [1]

MPEG-7 [2] is being developed by the Moving Pictures Expert Group (MPEG) [3], a working group of ISO/IEC. Unlike the preceding MPEG standards (MPEG-1, MPEG-2, MPEG-4) which have mainly addressed coded representation of audio-visual content, MPEG-7 focuses on representing information about the content, not the content itself.

The goal of the MPEG-7 standard, formally called the "Multimedia Content Description Interface", is to provide a rich set of standardized tools to describe multimedia content.

A single standard which can provide a simple, flexible, interoperable solution

to the problems of indexing, searching and retrieving multimedia resources will be extremely valuable and widely deployed. Resources described using such a standard will acquire enhanced value. Compliant hardware and software tools capable of efficiently generating and interpreting such standardized descriptions will be in great demand.

But will MPEG-7 be able to deliver such a standard -- one which satisfies its formidable goals and widely heterogeneous scope whilst concurrently providing simplicity, flexibility, interoperability and usability?

Objectives

MPEG-7 aims to standardize:

- a core set of **Descriptors** (Ds) that can be used to describe the various features of multimedia content;
- pre-defined structures of Descriptors and their relationships, called **Description Schemes** (DSs);
- a language to define Description Schemes and Descriptors, called the **Description Definition Language** (DDL);
- coded representations of descriptions to enable efficient storage and fast access.

MPEG-7 descriptions (a set of instantiated Description Schemes) will need to be linked to the content itself to allow fast and efficient searching for material of a user's interest. The descriptions may be physically located with the associated AV material, in the same data stream, on the same storage system, or the descriptions could be stored remotely. Hence mechanisms that can link the AV material to their MPEG-7 descriptions (and vice versa), regardless of where the content and its descriptions are located, are required.

Scope and Applications

MPEG-7 [4] is intended to describe audiovisual information regardless of storage, coding, display, transmission, medium, or technology. It will address a wide variety of media types including: still pictures, graphics, 3D models, audio, speech, video, and combinations of these (e.g., multimedia presentations). Examples of MPEG-7 data are an MPEG-4 stream, a video tape, a CD containing music, sound or speech, a picture printed on paper, or an interactive multimedia installation on the web.

MPEG-7 will address both retrieval from digital archives (pull applications) as well as filtering of streamed audiovisual broadcasts on the Internet (push applications). It will operate in both real-time and non real-time environments. A "real-time environment" in this context means that the description is generated at the same time as the content is being captured (e.g., smart cameras and scanners).

There are many applications and application domains which will potentially

benefit from the MPEG-7 standard. Examples of applications include:

- Digital libraries (image catalogue, speech archive);
- Broadcast media selection (radio channel, TV channel);
- Multimedia editing (personalised electronic news service, media authoring).

The potential applications cover a wide range of domains which include:

- Education;
- Journalism (e.g., searching speeches of a certain politician using his name, his voice or his face);
- Cultural services (museums, art galleries);
- Film, Video and Radio archives;
- Entertainment (e.g., video-on-demand, searching a game, karaoke);
- Investigation services (surveillance, human characteristics recognition, forensics);
- Geographical information systems;
- Remote sensing (cartography, ecology, natural resources management);
- Telemedicine and bio-medical applications.

Work Plan

Between October 1996 and October 1998, the scope, objectives and requirements for MPEG-7 were defined. The end of this stage was marked by an open Call for Proposals (CfP) in October 1998, which asked for submissions of relevant technologies fitting the requirements [5]. In answer to the CfP, some 60 parties submitted, in total, almost 400 proposals. The proposals were evaluated at the MPEG-7 Test and Evaluation Meeting in Lancaster in February 1999, according to their ability to satisfy the requirements. Certain proposals and elements of proposals were selected to be incorporated into the current collaborative phase.

Participants involved in making and evaluating submissions and the ongoing development of MPEG-7 include broadcasters, electronics manufacturers, content creators and managers, publishers and intellectual property rights managers, telecommunication service providers and academic researchers.

During the (current) collaborative phase, selected elements of various proposals are incorporated into a common model (the eXperimentation Model, or XM). The goal is to build the best possible model, which is in essence a draft of the standard. The XM is updated and improved in an iterative fashion until MPEG-7 reaches the Committee Draft (CD) stage, after several versions of the Working Draft. Improvements to the XM are made through Core Experiments (CEs). CEs are defined to test the existing tools against new contributions and proposals, within the framework of the XM, according to well-defined test conditions and criteria. Finally, those parts of the XM (or of the Working Draft) that correspond to the normative elements of MPEG-7 will be standardized. Table 1 illustrates the work plan.

Call For Proposals	October 1998

Evaluation	February 1998
First Version of Working Draft	December 1998
Committee Draft	October 2000
Final Committee Draft	February 2001
Draft International Standard	July 2001
International Standard	September 2001

Table 1. Scheduled Work Plan

Current State of the Descriptors

A Descriptor (D) defines the syntax and the semantics of one representation of a particular feature of audiovisual content. A feature is a distinctive characteristic of the data which is of significance to a user.

For example, the *color* of an image is a feature. Possible *Descriptors* corresponding to the *color* feature are: color histogram, RGB vector or a string. A Descriptor value is an instantiation of a Descriptor for a given data set. For example, RGB= (255, 255, 255), colorstring="red".

Table 2 illustrates some of the current descriptors which have been incorporated into the XM or are undergoing core experiments (CEs). They have been subdivided into Visual and Audio descriptors.

Type	Feature	Descriptors
Visual	Basic Structures	Grid layout
		Histogram
	Color	Color space
		Dominant color
		Color histogram
		Color quantization
	Texture	Spatial image intensity distribution
		Homogeneous texture
	Shape	Object bounding box
		Region-based shape
		Contour-based shape
		3D shape descriptor
	Motion	Camera motion
		Object motion trajectory
		Parametric object motion

		Motion activity
		Motion trajectory features e.g., speed, direction, acceleration
Audio	Speech Annotation	Lattice of words and phonemes plus metadata
	Timbre	Ratio of even to odd harmonics
		Harmonic attack coherence
	Melody	Melodic contour and rhythm

Table 2. Overview of Current Descriptors

Each descriptor is defined by normative and non-normative parts. The normative parts consist of the descriptor's syntax, semantics and binary representations of these. The optional, non-normative parts are the recommended extraction and similarity matching methods [6].

Many low-level features can be extracted from the content in fully automatic ways (e.g., color histogram). Recommended feature extraction algorithms are included in the non-normative parts of some descriptors. To allow for industry competition and to take advantage of expected improvements in technology, they are not a mandatory part of the standard. The same approach applies to similarity-based querying of descriptor values in which results are ranked in order of degree of similarity with the query. A recommended similarity matching method may be specified within a descriptor's non-normative component but it is not required for interoperability.

Some of the open issues regarding descriptors include:

- Is it possible to standardize certain descriptors (e.g., Timbre) without also standardizing the extraction and similarity matching methods?
- How can one compare the performance of descriptors with overlapping functionality in the CEs?
- How can one link procedural code (e.g., the extraction and similarity matching methods) to the description?
- How can one define complex composite descriptors such as parameterized arrays in the DDL?
- When does a composite descriptor become a description scheme?

Current State of the Description Schemes

A Description Scheme (DS) specifies the structure and semantics of the relationships between its components, which may be both Descriptors and Description Schemes.

The following concepts are used within the DS group to describe audiovisual content:

- Syntactic structure - the physical and logical structure of audiovisual content, e.g., structures based on temporal segments and/or spatial regions.
- Semantic structure - breakdown based on semantic meaning, e.g., structures based on temporal events and/or spatial objects.
- Syntactic-semantic links - the associations between syntactic elements and semantic elements.

The Generic Audiovisual DS [7] represents the integration of all of the DS proposals and submissions within a single DS. At the top level it consists of:

- A collection of Syntactic structure DSs, i.e., physical features such as segments, regions, color, texture, and motion are described here;
- A collection of Semantic structure DSs, i.e., semantic features such as objects, actors or events, e.g., "goal", "advertisement", "Madonna";
- Syntactic-semantic links DSs - which relate the syntactic elements to the semantic elements;
- Summary DS - this is used to enable browsing at different levels of granularity;
- MetaInfo DS - this contains descriptors carrying author or publisher-generated information, e.g., ContentDS, CreditsDS, CreationPurposeDS, RightsDS, PublicationDS, RightsDS;
- MediaInfo DS - this contains descriptors related to the storage media, e.g., file format, system, medium, colour, sound, length, duration, compression format;
- Model DS - this provides a way to describe the classification methods for audiovisual data or the correspondence between the current audiovisual content and other content through different models;

Figure 1 below illustrates the structure and content of the Generic Audiovisual DS.

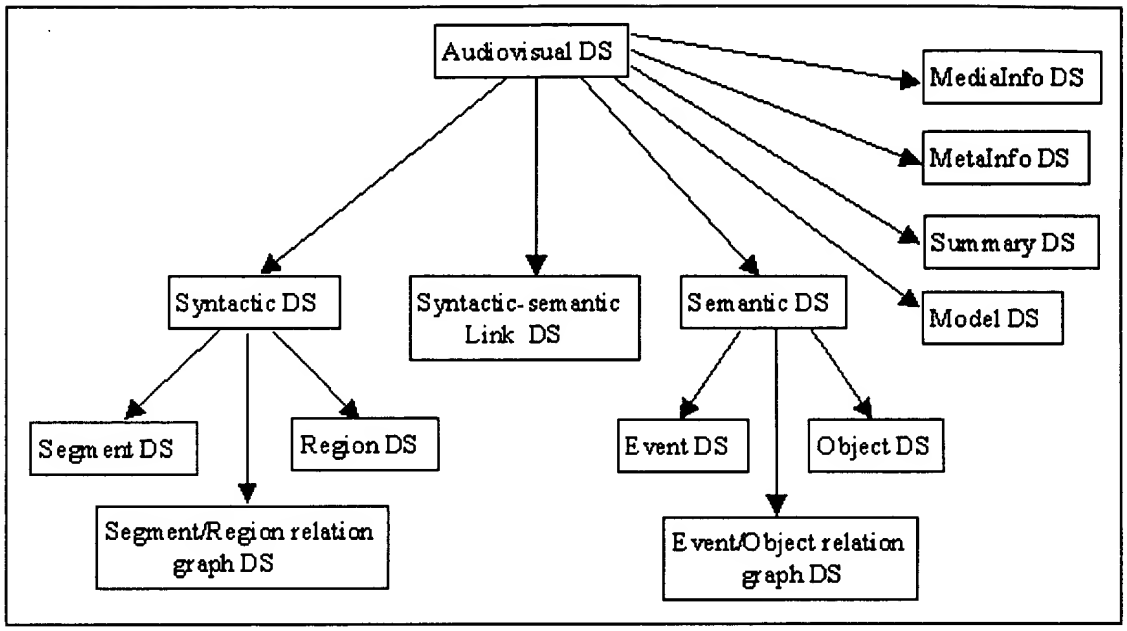


Figure 1. The Generic Audiovisual Description Scheme

One of the major problems with the DS work is the size and complexity of the Generic Audiovisual DS. There is a certain amount of redundancy and overlapping functionality between the different DS proposals which have been included. Some of the DS proposals which have been integrated are extremely complex and of dubious applicability. Unless a library of basic simple DSs is provided, many potential users who want simple bi-level multimedia metadata structures will find the MPEG-7 standard simply too bewildering or intimidating to use.

Current State of the Description Definition Language

The Description Definition Language (DDL) is the language that allows the creation of new Description Schemes and Descriptors. It also allows the extension and modification of existing Description Schemes.

The DDL has to be able to express spatial, temporal, structural, and conceptual relationships between the elements of a DS, and between DSs. It must provide a rich model for links and references between one or more descriptions and the data that it describes. It also has to be capable of validating descriptor data types, both primitive (integer, text, date, time) and composite (histograms, enumerated types). In addition, it must be platform and application independent and human- and machine-readable. The general consensus within MPEG-7 is that it should be based on XML syntax.

Of the ten DDL submissions which responded to the CfP in February, one was based on the Synchronized Multimedia Integration Language (SMIL), three were based on XML DTDs, three were based on XML DTDs with extensions such as data typing and inheritance, two were based on the Resource Description Framework (RDF) and one proposal was based on Open

Knowledge Base Connectivity (OKBC) [8].

After evaluating the DDL proposals, the recommendation was that -- although none of the proposals satisfied all of the requirements, the proposal from DSTC [9] provided the best starting point for further DDL development. However, it was also recommended that the DDL group should track the work of the W3C -- in particular, the XML Schema Working Group and the XLink, XPath and XPointer Working Groups.

In May this year, the XML Schema WG produced a 2-part working draft of the XML Schema language: XML Schema Part 1: Structures [10] and XML Schema Part 2 : Datatypes [11]. Discussions and preliminary encoding of the Generic Audiovisual DS led the DDL group to the decision to use XML Schema language as the basis for the DDL. However, certain reservations were raised at the Vancouver MPEG meeting in July concerning this approach. The major concerns were:

- MPEG-7's dependency on the output and time schedule of W3C XML Schema WG;
- Restricted access to internal documents associated with XML Schema development;
- The effect of W3C's copyright of XML Schema language on the ability to add MPEG-7-specific extensions.

As a result of these concerns, further discussions at the Vancouver meeting led to the decision to develop an MPEG-7-specific language in parallel with the XML Schema development being carried out within W3C [12]. A new grammar based on DSTC's proposal, but using MPEG-7 terminology (Description Schemes and Descriptors) and with modifications to ensure simple mapping to XML Schema, was recently developed. Based on this grammar, the following tasks are currently being performed:

- Specification of the BNF and an XML DTD for the new grammar;
- Specification of the validation mechanisms which must be provided by a parser;
- Development of a validating parser for this DDL

Relationship To Other Standards

MPEG-7 is aware of, and taking into account, the activities of a number of other standards groups during the development process.

For the archival descriptions, library (e.g., MARC, Z39.50) and archive (e.g., EBU/SMPTE, ISAD(G), EAD, Dublin Core, CEN/ISSS MMI) standards are being taken into account. Whilst for the streaming descriptions, the broadcast Electronic Programme Guides (EPGs) (e.g., DVB, ATSC) and web channels (Channel Definition Format (CDF)) standards are being considered. For the intellectual property and rights management descriptions, a liaison has been

formed with the INDECS project. The DDL group has been closely monitoring the work of the W3C's XML Schema Working Group and the XLink, XPath and XPointer Working Groups.

The MPEG-7 community is attempting to combine efforts with these groups through liaisons. This will hopefully maximize interoperability, prevent duplication of work and take advantage of work already done through the use of shared common ontologies, description schemes and languages. MPEG-7 hopes to act as a gateway or container for older established standards whilst at the same time providing a reference standard which can be used by proprietary multimedia applications or specific multimedia domains.

MPEG-7 Related Projects

There are undoubtedly a large number of MPEG-7-related projects being undertaken within commercial enterprises, particularly broadcasting and digital imaging companies, which involve the adoption of MPEG-7 conformance. The details of most of these projects are confidential. However, details are available for a number of collaborative government-funded research projects being undertaken, three of which are described below.

The HARMONY Project

HARMONY is a three-way International Digital Libraries Initiative project between Cornell University, the Distributed Systems Technology Centre and the University of Bristol's Institute for Learning and Research Technology. Its objective is to develop a framework to deal with the challenge of describing networked collections of highly complex and mixed-media digital objects. The research will draw together work on the RDF, XML, Dublin Core, MPEG-7 and INDECS standards, and will focus on the problem of allowing multiple communities of expertise (e.g., library, education, rights management) to define overlapping descriptive vocabularies for annotating multimedia content [13].

The DICEMAN Project

DICEMAN is an EC-funded project between Teltec Ireland DCU, CSELT (Italy), IBM (Germany), INA (France), IST (Portugal), KPN Research (Netherlands), Riverland (Britain) and UPC (Spain). Its broad objective is to provide an end-to-end chain for indexing, storage, search and trading of digital AV content. The technical work will focus on: MPEG-7 indexing through a COntent Provider's Application (COPA); the use of FIPA Agents to search and locate the best content; and support for electronic commerce and rights management [14].

The A4SM Project - A Framework for Distributed Digital Video Production

The A4SM project which is based at GMD's IPSI (Integrated Publication and Information Systems Institute) is currently researching the application of IT support to all stages of the video production process. The purpose is to seamlessly integrate an IT support framework into the production process, i.e., pre-production (e.g., script development, story boarding, etc.), production (e.g.,

collection of media-data by using an MPEG-2/7 camera, etc.), and the post-production (support of non-linear editing). In collaboration with tv-reporters, cameramen and editors they have designed an MPEG-7 camera in combination with a mobile annotation device for the reporter, and a mobile editing suite suitable for the generation of news-clips. [15]

Future Expectations

MPEG-7 is at a crucial stage of its development. In order to achieve wide-spread adoption as *the* standard for describing multimedia resources, MPEG-7 will have to resolve a number of formidable issues, including both high-level philosophical issues and low-level technical problems.

Some of the high level issues which need to be resolved include:

- Reconciliation of the opposing approaches of the various communities involved in MPEG-7 development which include:
    - the high-level semantic approach of the database/digital library community which typically believes that MPEG-7 needs only to provide standardised structure and linking mechanisms to the international community;
    - the low-level technical approach of the signal processing community which sees success in standardizing specific low level audiovisual features;
    - the free-spirited creative approach of the artistic content creators who don't like to be constrained or pigeon-holed by technocrats and their rigid rules, tools and best-practice guides.
- Striking the right balance between semantic and structural interoperability, media-specific and community-specific requirements and simplicity, extensibility and flexibility.
- Establishing and clarifying mutually-beneficial relationships between MPEG-7 and other existing standards bodies, e.g., W3C, Dublin Core, SMPTE.

Some of the low-level technical issues and problems which need to be resolved include:

- Integration of the Descriptor specifications within the Description Schemes;
- Refinement and clarification of the Description Schemes. Existing redundancies need to be removed before any new submissions are added;
- A decision must be made on the DDL, i.e., a choice between an MPEG-7 specific language or XML Schema language;
- Development of a validating parser for the chosen DDL;

- Provision of libraries of Descriptors and Description Schemes;
- Specification of (temporal, spatial, spatio-temporal, conceptual) links between descriptions and content;
- Enabling of links to procedural code -- extraction and similarity matching algorithms;
- Binary encoding of descriptions;
- Encoding of descriptions within streaming multimedia.

Assuming that the MPEG-7 participants do manage to overcome these obstacles, the success of MPEG-7 will then be dependent on the development and availability of hardware and software tools which can efficiently generate, store, search, retrieve and interpret MPEG-7 descriptions.

References

[1] Lesk, Michael, "Perspectives on DLI-2 - Growing the Field" *D-Lib Magazine*, July/August 1999. <http://www.dlib.org/dlib/july99/07lesk.html>

[2] MPEG-7 Home Page, <http://www.darmstadt.gmd.de/mobile/MPEG7/index.html>

[3] MPEG Home Page, <http://drogo.cselt.stet.it/mpeg/>

[4] MPEG-7 Requirements Group, "MPEG-7: Context, Objectives and Technical Roadmap, V.12", Vancouver, July 1999 ISO/IEC SC29/WG11 N2861

[5] MPEG-7 Requirements Group, "MPEG-7 Requirements Document V.9", Vancouver, July 1999 ISO/IEC SC29/WG11 N2859

[6] MPEG-7 Video Group, "MPEG-7 Visual Part of the eXperimentation Model Version 2.0", Vancouver, July 1999 ISO/IEC SC29/WG11 N2822

[7] MPEG-7 Description Scheme Group, "MPEG-7 Description Schemes (V0.5)", Vancouver, July 1999 ISO/IEC SC29/WG11 N2844

[8] MPEG7 Requirements Group, "Results of MPEG-7 Technology Proposal Evaluations and Recommendations", MPEG Seoul Meeting, March 1999 ISO/IEC SC29/WG11 N2730.

[9] Hunter J., DSTC , "A Proposal for an MPEG-7 DDL", P547, MPEG-7 AHG Test and Evaluation Meeting, Lancaster, February 1999

[10] XML Schema Part 1: Structures, W3C Working Draft, 6 May 1999 <http://www.w3.org/1999/05/06-xmlschema-1/>

[11] XML Schema Part 2: Datatypes, W3C Working Draft, 6 May 1999

<http://www.w3.org/1999/05/06-xmlschema-2/>

[12] MPEG-7 DDL Group, "MPEG-7 Description Definition Language Document V 1.0", Vancouver, July 1999 ISO/IEC SC29/WG11/ N2862

[13] The HARMONY Project <http://www.ilrt.bris.ac.uk/discovery/harmony/>

[14]The DICEMAN Project <http://www.teltec.dcu.ie/diceman/_vti_bin/shtml.exe/index.html/map>

[15] GMD IPSI - Integrated Publication and Information Systems Institute, Darmstadt <http://www.darmstadt.gmd.de/IPSI/movie.html>

Top | Contents
Search | Author Index | Title Index | Monthly Issues
Previous Story | Next story
Home | E-mail the Editor

DOI: 10.1045/september99-hunter

MPEG-7
Multimedia Content Description Interface

by
Jozsef Vass
2/2/98

Table of Contents

Overview

- Introduction
- History
- MPEG-7 Documents
- Schedule
- MPEG-7 Objectives
- Scope of MPEG-7
- MPEG-7 Requirements
- MPEG-7 Applications
- Possible Research Directions
- Case Studies -- QBIC, Columbia University

Introduction

- **Not** a compression standard

- MPEG-7 will be associated with the indexing of the available audio-visual (AV) information
- Before usage, information needs to be located -> *quickly* and *efficiently* search for various types of multimedia information
- Search for textual information exists (WWW search engines) but locating multimedia material is very difficult
- Information about content -> metadata
- Why MPEG-7? -> MPEG-5 would continue the series and MPEG-8 would give a binary pattern

---

History

Meeting	Date	Event
San Jose	February 1998	
Fribourg	October 1997	Refinements
Stockholm	July 1997	Documents
Bristol	April 1997	MPEG-7 Seminar

MPEG-7 Documents

- MPEG-7 Requirements Group, "MPEG-7": Context and objectives"
- MPEG-7 Requirements Group, "Second/third draft MPEG-7 requirements"
- MPEG-7 Requirements Group, "MPEG-7" applications document."

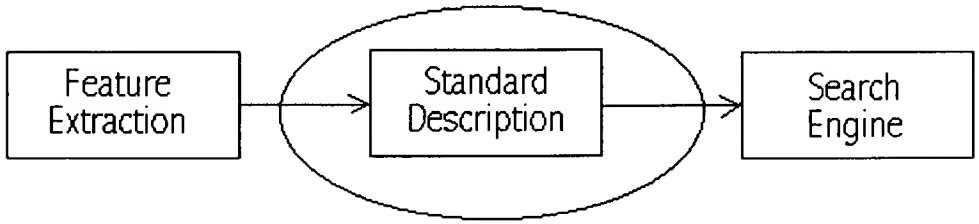
Schedule

Call for Proposals	November 1998
Working Draft	July 1999
Committee Draft	March 2000
Draft International Standard	July 2000
International Standard	November 2000

MPEG-7 Objectives

- Allow fast and efficient searching for multimedia material of user's interest
- Specify a standard set of *description schemes* and *descriptors* to describe various types of multimedia information
- Specify encoding of description schemes and descriptors
- Descriptors are associated with the content itself
- Descriptors must be meaningful in the context of application
- MPEG-7 data can be located anywhere
- Descriptors are independent of the encoding of the content
- Any type of AV material can be retrieved by any type of query material

Scope of MPEG-7



Scope of MPEG-7

- MPEG-7 is general
- Only the descriptor is standardized

MPEG-7 Requirements

Requirements are divided into 1) common audio and visual requirements, 2) visual requirements, and 3) audio requirements

- Query Classes: MPEG-7 shall support audio and visual descriptors that allows different types of queries:
  -- Textual Description Classes: keywords, title, author, structure, etc.
  -- Visual Description Classes: color, texture, shape, volume, etc.

-- Audio Description Classes: frequency contour, timbre, harmony, textual content, etc.
- Content-Based Retrieval: Support effective and efficient retrieval of multimedia based on their content
- Types of Features: Various types of features ranging from low level to high level features
- Feature Hierarchy and Scalability: Coarse-to-fine refinement of features
- Distributed Databases: Support of simultaneous and transparent retrieval
- Robustness to Information Errors and Loss
- Ability to export existing databases into MPEG-7 format
- Copyright Information

MPEG-7 Applications

Application domains:

- Education
- Journalism
- Entertainment
- Tourism
- Medical Applications
- Archives (Film, Video, and Audio)
- Shopping
- Social

Possible applications:

- Intelligent Multimedia Presentation: By using knowledge of context, application, location of user the system can adapt to the user (similar to Web personalizing)
- Storage and Retrieval of Video Databases: Enormous interest to search end retrieve from existing huge radio and TV archives
- Teleshopping: Humans remember better to visual information than textual. Internet catalogs could be searched much easier by using visual query.
- Entertainment: For example Karaoke user interface can be made more convenient by query by example (sing part of song)
- Medical: Physicians can search huge medical archives for similar symptoms by query by example

---

Research Directions

- Combination of low level information (extracted automatically) with high level description (mostly obtained manually)
- Data structures
- How MPEG-4's AV objects can be extended to support the requirements of MPEG-7?
- Language independence

Case Study - QBIC

- QBIC: Query by Image Content
- Computing image features and retrieval based on these features
- Can be extended to video retrieval by using key frames or mosaics
- More information: http://wwwqbic.almaden.ibm.com/

Case Study - Finding Images/Video Archives

- Image-Based Visual Query Project
- WWW-based
- Retrieval is based on features (motion, spatial, visual, syntactic or semantic characteristics of objects)
- Operating in compressed domain (MPEG, not progressive) \item More information: http://www.ctr.columbia.edu/~sfchang

---


CECS Multimedia Communications and Visualization Laboratory

The integration of Metadata from production to consumer

Peter Mulder
Dutch Broadcast Facilities Company N.V. (NOB)

The newly-formed MPEG-7 Ad-hoc Group on Integration is currently dedicating itself to the task of integrating metadata as approached by the SMPTE (for the professional TV Production domain) with the metadata approach chosen by the MPEG-7 community. The SMPTE approach is based on a dictionary and binary coding, and is intended specifically for machine control and fast real-time applications. The MPEG-7 approach is based on standard XML and is human readable.

For professional use during the more technical phases of production and post production, the SMPTE approach can be well suited while, in the domain of consumer set-top boxes, the most promising interface is XML-based.

Both approaches have value in their own right, each with distinct advantages at specific points in the content production and delivery processes. For this reason alone, it is worth the effort of trying to harmonize the two approaches. It would be of great benefit to broadcasters if production metadata and consumer services were to connect together seamlessly – without human intervention in the form of Metadata Editors in the transmission multiplex area.

Introduction

In many different groups, people are working on metadata and its standardization. Some of these groups have a special interest in the broadcasting world, in its widest sense. Others are interested only in the highly-specialized professional production environment.

As a result of the output from the EBU/SMPTE Task Force's work on *the harmonization of standards for the exchange of programme material as bitstreams (TFHS)* [1], standards development has proceeded within the SMPTE's Technical Committee, particularly

group W25, and also within other organization's committees such as the EBU P/Meta group [2]. While the SMPTE is working in the broad domain, including everything related to programme-making in a broadcast production environment, P/Meta is focusing on the area of Business-to-Business and System-to-System exchanges between Broadcasting or Production organizations, although they recently became involved also in the Broadcaster-to-Consumer domain.

In parallel with the start of this process, the MPEG community also realized that there was a pressing need for the development of a comprehensive MPEG metadata standard. This work started in July 1997 as MPEG-7. The work in MPEG-7 has, until recently, been largely driven by Academia and the Telco companies. Focus has very much been on web-based applications and annotation tools for audio and video material. The work has concentrated on the development of tools for describing concepts and content, although there also has been an interest in video editing.

However, MPEG-7 did not limit its scope to broadcasting, but extended into many other domains also; for example, medicine, physics and many other applications involving the description of audio-visual content. About one year ago, the broadcast community started to actively participate in this work and, of course, looked to the metadata work that was being done in the area of broadcast-related applications by a small group of specialists – largely the same group that was also participating extensively in the SMPTE metadata work. This group came to the conclusion that it was essential to harmonize at least the MPEG-7 and SMPTE standards, in view of the potential benefit and great opportunities opened up if these two work areas were to complement each other and could be made to map closely to each other – and the even greater benefits that would arise if related work, such as Dublin Core, the FIAT minimum data list and the work of the INDECS project, could also be incorporated. From that perspective, the MPEG-7 Integration Group was set up to formulate proposals for a framework that would allow the interoperability of metadata systems targeted at production, knowledge management, post production, archival repository, distribution, publication, and the exchange of audio-visual material both between businesses and between businesses and consumers.

This group quickly identified the need for a concise and consistent dictionary of terms and definitions within the various schemes, as well as the need for concise mappings between the MPEG-7 and SMPTE etc. work: indeed, a common dictionary would be an ideal outcome although this is likely to be extremely difficult to achieve. If MPEG-7 can be made to fit on top of, or act as an extension to library and broadcast production metadata, then there will be an opportunity, with little extra cost, to achieve intelligent navigation, multimedia handling and the exchange of descriptive data (metadata) between content providers (producers), professional and other users (e.g. broadcasters), and final users (consumers) – as well as unlocking the potential of effective knowledge management from data-rich libraries of stored material.

The MPEG-7 Integration Group has identified a validation process to enable the migration towards a common standardized metadata layer, underpinning all the business processes from the earliest conceptual stages of production, through the full range of multimedia production and post production processes, right into the home platform.

This process will enable the application of Information Science techniques to unlock the full potential of archival storage in such a way that existing content can be retrieved for a multitude of purposes.

As things stand, the professional production of multimedia content (including its metadata) is likely to be largely based on standards each written for a specific fragment of the end-to-end industry process. For instance, in the production and post production phases of standardization, it is likely to revolve around SMPTE standards such as the SMPTE Metadata Dictionary (SMPTE 335M and RP210), the Unique Material Identifier (UMID) (SMPTE 305M) and Key Length Value coding (SMPTE 336M). For librarianship purposes, the Dublin Core or the FIAT minimum data set is widely referenced while, in the conceptual and consumer domains, MPEG-7 is a very promising newcomer. However, it is essential that, as standards develop, applications based on any one of these standards should seamlessly integrate with applications further down the chain which are based on another standard. In particular, it is essential to ensure integration with consumer products using, for instance, XML so that metadata is passed transparently through the chain from the conception of an idea by the producer to consumption by a viewer or listener – without error or human intervention.

In the MPEG-7 Ad-hoc Group on Integration, the work has just been started to build the first version of an MPEG-7 dictionary. The group has recently taken the first steps in this process and, as a convenient starting point, will focus on filtering off Descriptors from the MPEG-7 Descriptor Schemes and, hence, produce a flat list of the Descriptors with their definitions and data-type. The next step will be to study if the MPEG-7 schema and SMPTE Dictionary world views can be reconciled and, currently, this work is in its very earliest stage. Once this preparation of the dictionary of terms and definitions has progressed to first draft stage, it will be possible to study the potential for the integration of the two. This must be done in such a way that precise definitions of metadata "elements" can be compared between the SMPTE and MPEG dictionaries, which makes the work painstaking and not a little challenging!

Initial work has, significantly, already revealed that the discipline of this dictionary approach is in any case essential and of vital importance in keeping the MPEG-7 schemes compliant within themselves. It has also demonstrated its suitability as a basis for interoperability with other metadata systems.

Abbreviations

FIAT	*Fédération Internationale des Archives de Télévision* (IFTA in English)
KLV	(SMPTE) Key Length Value
MPEG	Moving Picture Experts Group
NOB	Dutch Broadcast Facilities Company N.V.
SMPTE	Society of Motion Picture and Television Engineers (USA)
TVA	TV-Anytime
UMID	(SMPTE) Unique Material Identifier
XML	Extensible markup language

Although the MPEG-7 schemes enable a very rich mix of conceptual descriptions, there is currently a lack of an MPEG-7 transport and coding mechanism and this issue is under consideration at the moment in the Binary Format ad hoc group. The SMPTE has recently standardized Key Length Value (KLV) coding for transporting metadata within professional production technical systems. This coding protocol is specifically intended for transporting metadata associated with multimedia files and can be very bit-efficient. The study group in MPEG considers this way of coding as one of the possibilities for a binary transport of MPEG-7 metadata. One of the essential tools to be developed is one that will translate KLV into XML and vice-versa. In order to do this in a reliable way requires a common understanding of the basic elements used to build the applications and schemes. In particular, the integration of web-based consumer and professional user applications, alongside other professional production and post production tools in a professional production environment, will require complete interoperability in applications such as the searching of professional databases.

The broader picture

From the viewpoint of end-users, integration is essential since the industry literally cannot afford the costs of living with competing and possibly incompatible schemas in different parts of the production-to-consumer chain. Integration will also enable the elimination of unnecessary differences between schemas, and the minimizing of translation processes at interfaces. While in theory, it may be possible to work with different standards in each domain (production, post production, content management and distribution) – provided there is sufficient compatibility to allow automatic translation at the interfaces – this is undesirable. Translation is inefficient and unreliable (or lossy) without intervention, and is likely to be expensive to resource.

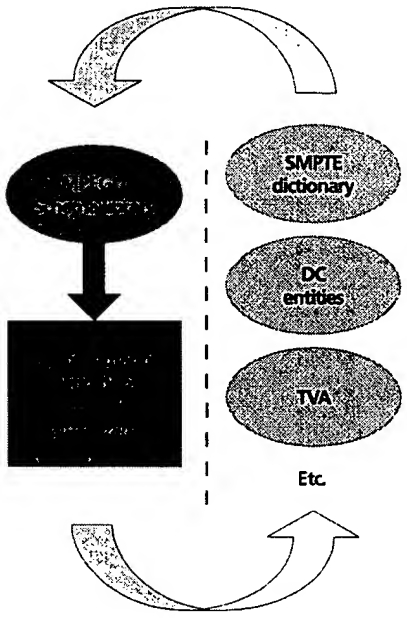


Figure 1
The process for metadata integration

Clearly, the better alternative is to have a common structure and compatible domain-specific vocabularies with a single or federated public registry of the vocabularies.

Hence, the next stage in the integration process will be to study the integration needs beyond those of the SMPTE and EBU or Dublin Core, and into areas such as NewsML.

Fig. 1 outlines the process for integration:

The process on the left side represents the MPEG-7 environment and that on the right, the "others".

Peter Mulder *is currently a Consultant of Technology and Strategy development at NOB Department of Technical Services. He joined the company in February 1964 and later (from 1972 till 1989) worked at the Research and Development Department. From 1989 until 1994, he was Production Manager at NOB Design (graphics department). In 1994 he became involved in setting up of the then new business unit "NOB-Interactive", a unit that is still engaged in the development of new (multi)media and distribution possibilities and also with developing new technical infrastructures for performer-driven animation and for virtual studio systems.*

Mr Mulder is active in many related international standards committees such as the SMPTE engineering committees, EBU/P-Meta, MPEG-7 and the TV-Anytime forum. During 1998, he was elected SMPTE Governor for the Europe, Middle East and Africa region and later re-elected for the 2000-2001 term. His main interest at the moment is in metadata and in both its use and exchange between the professional and consumer domains – "From shooting to the Personal Disk Recorder".

When a new MPEG-7 schema is proposed, it will be necessary to submit a contribution to the MPEG-7 MDS group and to agree the new application schema. This will be examined within the MPEG community and, if it fits in with the overall Standard structure, the scheme can be accepted.

Similarly, proposals in the other domains will be submitted in that domain.

Since the integration work has ensured reliable mappings between the standards, the mapping process for new submissions can then be completed without difficulty.

Closing remarks

In both the SMPTE and MPEG-7, the metadata groups are very small with only a few representatives from the broadcasting community. The issue is, however, vitally important to this community also: any interested broadcasters are invited to, indeed should, participate in this metadata work.

Bibliography

[1] EBU / SMPTE Task Force: **Final Report: Analyses and Results** *(2554 KB)* "Special Supplement 1998" to EBU Technical Review.

[2] R. Hopper: **EBU Project Group P/META – Metadata Exchange Standards** EBU Technical Review No. 284, September 2000.